University of Alberta

Library Release Form

Name of Author: Andrea Gale Kiceniuk

Title of Thesis: Development of Capillary Electrophoresis/Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Proteome Analysis

Degree: Master of Science

Year this Degree Granted: 2001

I have no more faith in men of science being infallible than I have in men of God being infallible, principally on account of them being men.

-Noël Coward

University of Alberta

Development of Capillary Electrophoresis/Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Proteome Analysis

by

Andrea Gale Kiceniuk

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

Department of Chemistry

Edmonton, Alberta
Fall 2001

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Development of Capillary Electrophoresis/Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Proteome Analysis submitted by Andrea Gale Kiceniuk in partial fulfillment of the requirements for the degree of Master of Science.

For my husband,
Andrew

Abstract

A method of combining capillary electrophoresis (CE) using surfactant-modified capillaries with matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) is described for proteome analysis. Microliter volume fraction collection is performed on CE to facilitate off-line coupling with MALDI MS. It is shown that the use of dilute acid instead of electrophoretic buffer for fraction collection lowers the detection limit of MALDI MS. The fraction collection method also allows reactions such as enzymatic digestions to be carried out on fractioned samples for protein identification. Several digestion techniques for peptide mapping of CE-fractioned proteins are evaluated. CE/MALDI MS is then applied to the analysis of a complex protein mixture from *Escherichia coli* cell lysates. The use of liquid chromatography prior to CE is found to reduce ion suppression during MALDI analysis and to allow the detection of more proteins. Several challenges encountered when analyzing complex proteomes, and techniques to overcome them, are discussed.

Acknowledgements

I would like to thank my supervisor, Dr. Liang Li, for his support and guidance over the past four years. His enthusiasm and dedication have been an inspiration. I would never have reached this point without his help.

I would also like to thank the other members of my examining committee, Dr. Chuck Lucy and Dr. Peter Sporns, for their careful review of my thesis. In addition, I thank Dr. Paul D'Agostino of the Defence Research Establishment at Suffield for helpful discussion during the course of this work.

Many thanks to all the other group members, past and present, with whom I have worked. They have been wonderful people to know and I have made many good friends. My special thanks go to Dr. Bernd Keller for his help with MS samples and nanochem station work, to Ms. Zhengping Wang for her help with the *E. coli* samples, and especially to Dr. Ken Yeung for all his CE expertise and mentoring.

I thank all the personnel in the Chemistry Department for taking care of the administrative and technical details that helped me through my graduate program. I also thank the Chemistry Department and the University of Alberta for providing excellent facilities and a stimulating learning environment.

I thank NSERC for PGS A and PGS B scholarships and the University of Alberta for accompanying Walter H. Johns Graduate Fellowships during the four years of my program. I thank the Chemistry Department and the University of Alberta for graduate teaching assistantships and graduate research assistantships. Finally, I thank the Department of National Defence of Canada for research funding.

Most of all, I thank my husband Andrew for his assistance while preparing this thesis, and for his encouragement, love and support throughout my studies.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

amol	attomole (1amol = 10^{-18} mol)
CE	capillary electrophoresis
CIEF	capillary isoelectric focusing
CZE	capillary zone electrophoresis
DDAB	didodecyldimethylammonium bromide
Da	dalton, 1 Da = 1 u (atomic mass standard)
EOF	electroosmotic flow
E. coli	*Escherichia coli*
fmol	femtomole (1fmol = 10^{-15} mol)
HCCA	α-cyano-4-hydroxycinnamic acid
I.D.	inner diameter
kPa	kilopascal (1 kPa = 10^{3} Pa)
LC	liquid chromatography
m/z	mass-to-charge ratio
μg	microgram (1 μg = 10^{-6} g)
μL	microliter (1 μL = 10^{-6} L)
μm	micrometer (1 μm = 10^{-6} m)
μM	micromolar (1 μM = 10^{-6} M)
MALDI	matrix-assisted laser desorption/ionization
MS	mass spectrometry
nL	nanoliter (1 nL = 10^{-9} L)
PAGE	polyacrylamide gel electrophoresis
PEG	poly(ethylene glycol)
TFA	trifluoroacetic acid
TLF	time-lag focusing
TOF	time-of-flight
v/v	volume-to-volume ratio

Chapter 1

An Introduction to Capillary Electrophoresis and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Protein Identification

1.1 Proteomics: Strategies for Protein Identification

Gene sequencing techniques, including the use of capillary electrophoresis[1-4], have advanced rapidly in recent years. This has resulted in a "working draft" of the human genome (http://www.ornl.gov/TechResources/Human_Genome/home.html) and the complete sequencing of the genomes of over twenty simpler organisms[5]. These advances in genomics have naturally led into the study of the proteome, or the complete set of proteins expressed by the genome of a cell under a given set of conditions. Ultimately, researchers would like to understand the functions and interactions of all the proteins, as well as the factors controlling their expression. The first step, however is to identify and characterize the proteins.

Mass spectrometry (MS) has become the principle technique used for protein identification in proteome analysis[5-8], offering faster analysis times and higher sensitivity than the alternative Edman sequencing technique. The initial step in identifying a protein is often the use of MS to determine its molecular mass. However, additional information is needed to accurately identify an unknown protein. There are two main mass spectrometric methods used to acquire this information: peptide mapping (sometimes called peptide mass fingerprinting) and partial peptide sequencing by tandem mass spectrometry. For peptide mapping, the protein is chemically or enzymatically digested and the resulting peptide mixture is analysed by MS. The protein is then identified by matching the peptide masses against libraries of theoretical protein digests. Several digestion methods can be used on the same protein sample to increase the confidence of the identification. For peptide sequencing by tandem mass spectrometry, the protein is also digested and introduced into a mass spectrometer. A mass spectrum of the peptide

mixture is acquired, and then the peptides are individually isolated and fragmented by collision-induced dissociation (CID). The resulting fragment ions are compared against theoretical fragmentation patterns of peptides from theoretical digests of proteins in a database using a computer program such as SEQUEST (distributed by Thermo Finnigan). The program returns a list of possible protein matches with numerical scores to help determine the best match.

Although MS is very powerful, the protein mixtures encountered in proteomic research are usually too complex to be effectively analysed by MS alone. The more-abundant components will significantly suppress the signals from the less-abundant components. For example, a typical mass spectrum of proteins from a cell extract is dominated by a few abundant proteins, even though hundreds of proteins are present at concentrations above the MS detection limit. The signals of many proteins are lost as a result of ion suppression. To avoid this suppression, MS can be combined with a separation technique to isolate the various components in the mixture prior to mass spectral measurements. Two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) is widely used for proteome display, but it is a time-consuming and laborious technique. Liquid separations such as liquid chromatography (LC) and capillary electrophoresis (CE) are becoming more prevalent as an alternative to 2-D PAGE for interfacing with MS. This work investigates the off-line coupling of CE with matrix-assisted laser desorption/ionization (MALDI) MS for protein analysis.

1.2 Capillary Electrophoresis

1.2.1 Basic Principles of CE Electrophoresis is the separation of charged analytes based on their differential migration rates in an applied electric field. The higher the applied voltage, the faster the analytes will migrate, so high operating voltages are desirable for faster analysis times. However, resistive heating, called Joule heating, becomes worse as the voltage is increased. This heating causes convection in the separation medium and distorts the analyte bands, decreasing resolution. To counteract the effects of Joule heating, electrophoresis has classically been performed at low voltages using a solid support such as paper or polyacrylamide gel. However, Joule

heating can also be minimized by performing electrophoresis in a small capillary, i.e. 'capillary electrophoresis'. The high surface area-to-volume ratio of small capillaries allows heat to be quickly dissipated. As a result, high separation voltages can be used, which leads to faster analysis times.

Capillary zone electrophoresis (CZE) refers to CE that is performed in a buffered solution without additives such as surfactants or ampholytes to affect the separation mechanism. CZE was first demonstrated by Hjerten[9] in 1967 using 3 mm tubes. In the following decades, the technique was refined by the use of capillaries with smaller inner diameters (I.D.s) to improve resolution and efficiency[10-12]. Today, CE is usually performed in 10-100 μm I.D. fused silica capillaries covered with a polyimide coating to increase their flexibility for easier handling. A typical CZE system is shown in Figure 1.1. The ends of the capillary are immersed in buffer reservoirs that are connected to a high voltage (HV) power supply by electrodes. Near the outlet end of the capillary, a small window is burned in the polyimide coating to allow for optical detection of analytes. In this work, a diode array detector (DAD) was used for absorbance detection. Data from the detector is collected by data acquisition software and stored in a computer. Most CE systems also have an external cooling system to help dissipate heat from the capillary.

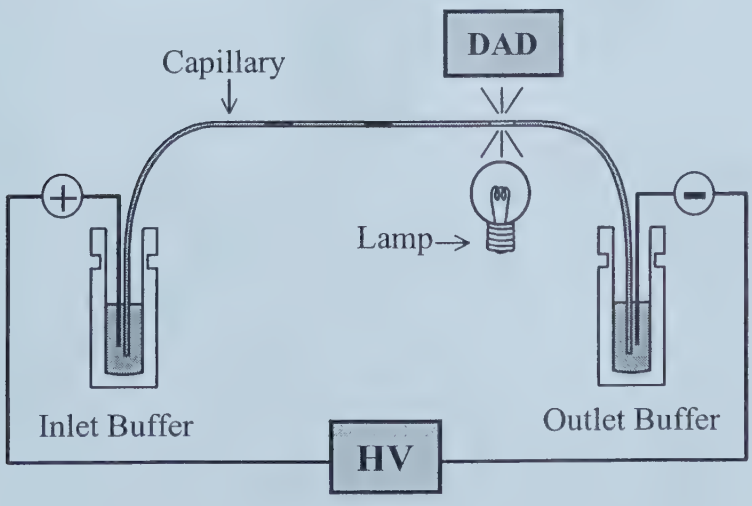

Figure 1.1 Schematic diagram of a typical CE instrument.

To perform a separation, the capillary is filled with buffer solution by applying pressure at the inlet reservoir. The inlet reservoir is then exchanged for a sample reservoir, and a small sample plug is introduced into the capillary using applied pressureor voltage. After replacing the buffer reservoir, a voltage is applied across the capillary and the analytes separate according to their electrophoretic mobilities.

Movement of the bulk solution through the capillary in CZE separations is not generated by a pump, but rather by electroosmotic flow (EOF). EOF is generated as a result of the surface charge on the capillary wall. Silanol groups (-SiOH) on the fused silica surface become ionized ($-SiO^-$) between pH 3 and 8. In order to maintain electroneutrality at the anionic surface, an electric double-layer is formed at the solid-liquid interface[13] (Figure 1.2). Some cations in the CE buffer solution adsorb on the anionically-charged wall, forming an immobilized compact layer called the Stern layer. The remaining cationic counterions are distributed in a diffuse layer. When an electric

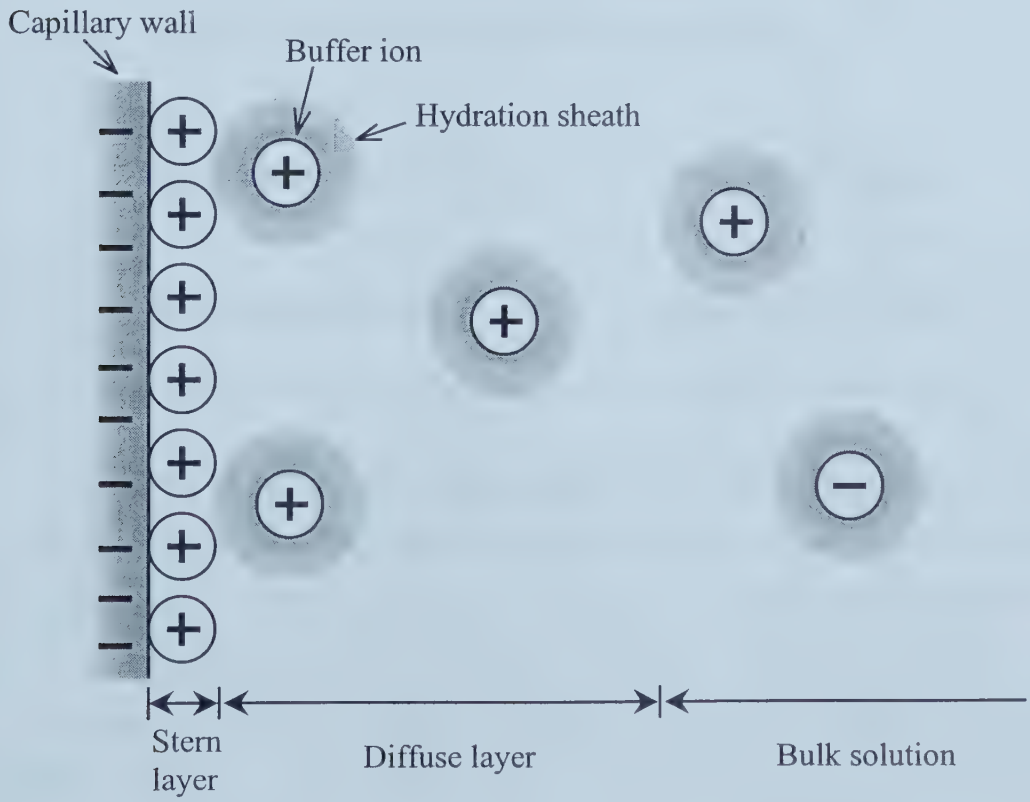

Figure 1.2 The electrostatic double-layer generated inside a fused silica capillary.

field is applied, the cations in the diffuse layer move towards the cathode. Since the cations are hydrated, they drag the solvent along with them. This results in a movement of the bulk solution called electroosmotic flow.

When a coating is applied to the inner wall of a capillary, the silanols in the fused silica are masked and the wall will adopt the charge of the coating. In this way, capillary coatings may alter the direction and magnitude of the EOF. In this work, a cationic capillary coating is used. As a result, the diffuse layer is populated by anions rather than cations and the EOF is reversed, moving towards the anode.

The EOF velocity increases radially from zero at the Stern layer to its full magnitude at the boundary between the diffuse layer and bulk solution. The thickness of the diffuse layer, however, is negligible compared to the inner diameter of the capillary, so the EOF is essentially constant across the capillary. In addition, the driving force of the EOF is distributed evenly all along the capillary, so the flow is nearly uniform throughout[14]. These factors lead to a flat plug-like flow profile (Figure 1.3a) as opposed to the parabolic flow profile generated by a pump (Figure 1.3b). This flat flow profile reduces bandbroadening, resulting in higher separation efficiencies.

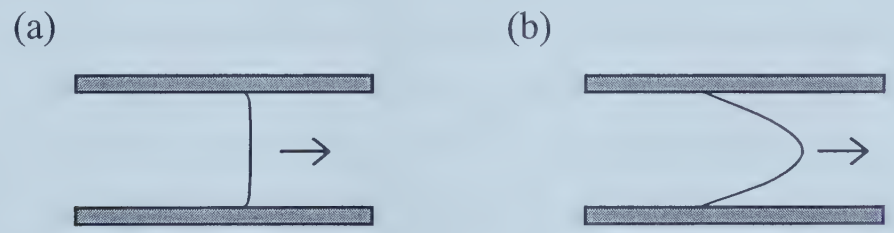

Figure 1.3 Flow profiles produced by a) electroosmotic flow and b) laminar flow.

When placed in an electric field, anions will migrate towards the anode and cations towards the cathode. This action is an intrinsic quality of charged species and is called electrophoretic mobility (μ_e). The electrophoretic mobility of a molecule is a function of the electric force that the molecule experiences and the frictional drag in the separation medium. Mobilities can be estimated from Debye-Huckel-Henry theory using the equation:

$$\mu_e = \frac{q}{6\pi\eta r} \qquad \text{(Equation 1.1)}$$

where q is the net charge of the charged molecule, r is its radius and η is the viscosity of the separation medium. While this equation was derived for spherical species, it still provides a reasonable approximation for real molecules. The velocity (υ) of a migrating molecule is in turn given by:

$$\upsilon = \mu_e E \qquad \text{(Equation 1.2)}$$

where E is the applied electric field (V/cm). From Equations 1.1 and 1.2, it can be seen that a molecule's mobility and therefore its migration velocity is determined by its charge and size. In general, smaller molecules migrate faster than larger molecules, and highly-charged molecules migrate faster than weakly-charged molecules. In this manner, different molecules will separate from each other during electrophoresis.

When looking at the overall behaviour of analytes during CZE separation, both the EOF and an analyte's electrophoretic mobility must be taken into account. The observed or apparent mobility (μ_a) of a molecule is the sum of its electrophoretic mobility (μ_e) and the electroosmotic mobility (μ_{eo}):

$$\mu_a = \mu_e + \mu_{eo} \qquad \text{(Equation 1.3)}$$

Figure 1.4 shows the resulting apparent motion of positively, negatively and neutrally charged molecules. Assuming that a cationically-coated capillary is used, the EOF will be anodic. Anions will migrate with the EOF, cations will migrate counter-EOF and neutrals will be carried along with the EOF. If reversed polarity is used, i.e. if the detector is at the anodic end of the capillary, the analyte migration order will be anions,

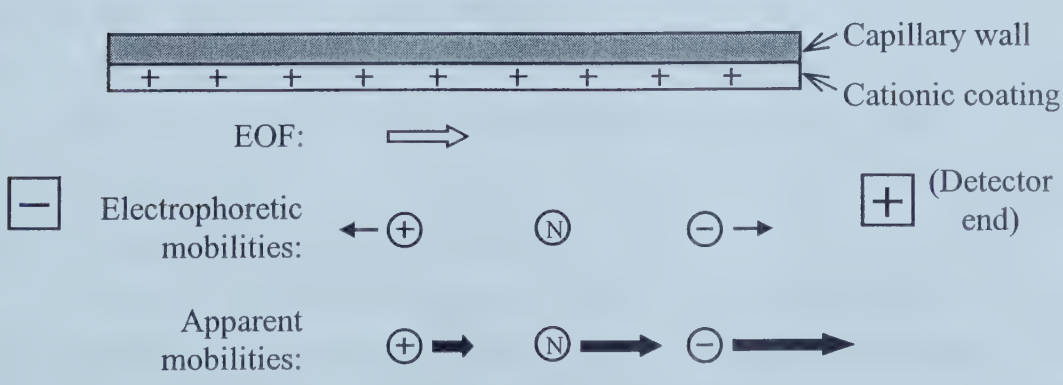

Figure 1.4 Apparent mobilities and migration order of analytes during electrophoresis. Analytes are labelled + for cations, N for neutrals and – for anions.

neutrals and then cations. Note that although neutral molecules become separated from charged molecules, they are not separated from one another because they do not have electrophoretic mobilities themselves.

1.2.2 Coupling CE with MALDI MS There are several reports of combining liquid separation with MALDI MS. In addition to the continuous flow probe method for LC-MALDI coupling[15-17], other on-line approaches have also been described[18]. However, on-line LC or CE/MALDI MS has not been widely used for real-world applications due to current technical challenges, which lead to less than desirable performance for peptide and protein analysis[19]. Alternatively, off-line coupling between LC and MALDI MS can be performed by fraction collection. The offline coupling of MALDI-MS with CE, however, is less straightforward as the sample size in CE can be as low as a few nanoliters. In addition, electrical contact must be maintained during the electrophoretic separation. Methods reported for coupling CE and MALDI MS generally involve either direct sample deposition from the CE capillary onto a MALDI target[20] or CE fraction collection followed by subsequent MALDI analysis[21].

Direct sample deposition for off-line coupling of CE with MALDI-MS results in excellent sensitivity. The sample is deposited onto targets pre-coated with MALDI matrix. All of the sample injected in CE is deposited with minimum handling. Sample deposition can be performed on a stationary MALDI plate in which the sample is spotted as droplets[22-24]. Alternatively, sample can be deposited as a trace on a moving MALDI target, such as a tape or a wheel[25-28]. The lowest detection limits reported for peptides and proteins using direct sample deposition are 50 attomoles and 2 femtomoles, respectively[28]. These numbers refer to the quantities injected into CE, and they are very close to the detection limits of the current MALDI mass spectrometers. While direct sample deposition offers low sensitivity, it does require special instrument alteration, which is not easily accessible by other users at present.

Fraction collection does not require special instrument alteration, and it can be readily performed with commercial instruments. In addition, the fractionated samples can potentially be divided into smaller portions for down-stream chemical or enzymatic processing. As discussed in Section 1.1, the latter is crucial for protein identification and

analysis of post-translational modifications. Multiple chemical or enzymatic reactions are necessary for achieving high amino acid coverage of a protein, thus increasing the confidence of identification or providing structural information on the modification sites. For these reasons, the micro fraction collection approach is chosen in this work.

To perform fraction collection in CE, the electrophoretic separation is interrupted just before an analyte band reaches the outlet of the capillary. The analyte band is then collected either hydrodynamically or electrokinetically[21]. When hydrodynamic fractionation is used, a low pressure is applied at the inlet to push the sample band off the capillary into a micro vial. However, significant band broadening is caused by the pressure-induced laminar flow during fraction collection. More importantly, the mobilities of the analytes remaining in the capillary are interrupted, and thus subsequent fraction collection is not accurate. Electrokinetic fraction collection is therefore preferred. When electrokinetic collection is used, the analyte band is mobilized with an applied voltage. This maintains high separation efficiency, but salts from the collection electrolyte can be a problem for subsequent MALDI analysis. The optimization of electrokinetic fraction collection for CE/MALDI MS analysis is discussed further in Chapter 2.

1.3 MALDI-TOF MS

1.3.1 The MALDI Process The use of matrix-assisted laser desorption as an ionization technique was first reported in 1987 by two independent groups[29,30]. Most MALDI techniques developed since that time are variations of the method used by Karas *et al.*[30,31] which employed solid matrices. A schematic diagram of the MALDI process is shown in Figure 1.5. To prepare a sample for MALDI analysis, the sample is mixed with a solution containing the matrix and then deposited on a metal target. As the mixture dries, the analyte becomes incorporated in the crystal structure of the matrix. The target is then introduced into the high-vacuum (~10^{-6} Torr) source region of the mass spectrometer and a laser pulse is fired at the sample spot. In this work, UV-MALDI was performed with a nitrogen laser (337 nm). The laser pulse generates a MALDI plume, which can be

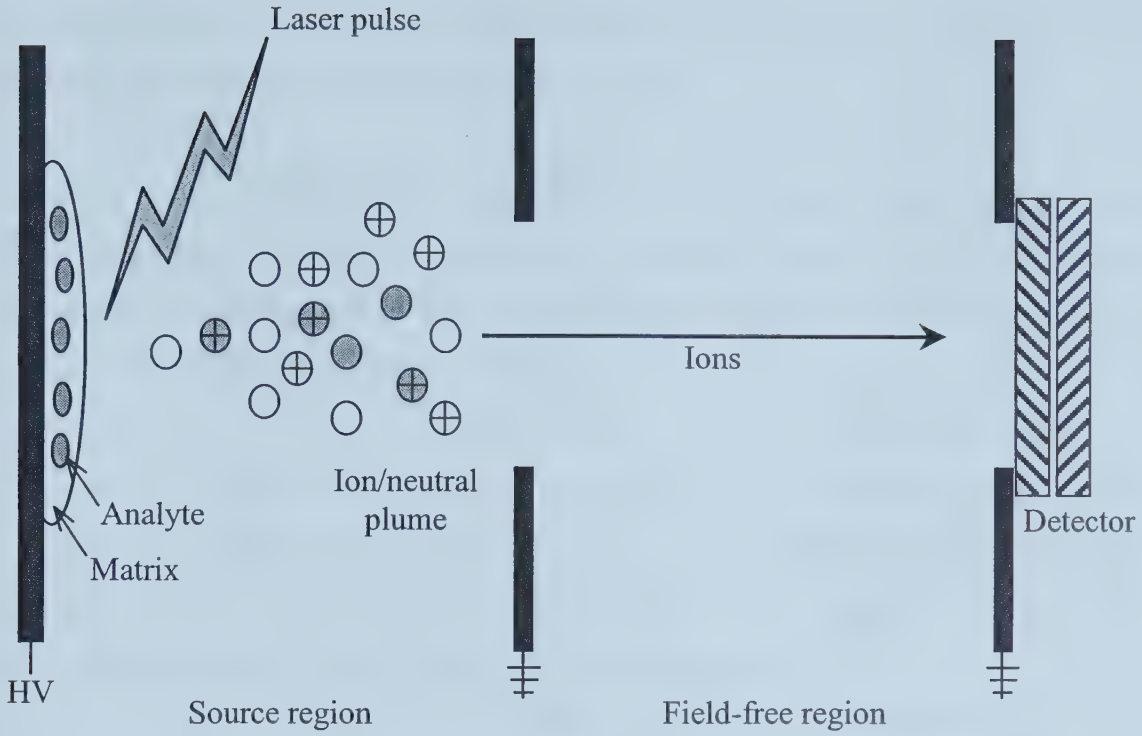

Figure 1.5 Schematic diagram of a linear MALDI-TOF mass spectrometer.

described as a rapid solid-to-gas phase transition for the analyte and matrix molecules. The plume contains a high density of neutral matrix molecules as well as ionic, radical and electronically excited species. The exact reactions that occur in the plume to create matrix and analyte ions are still in dispute, with several ionization processes and models proposed[32,33]. It is generally believed that primary ions are formed through processes such as multiphoton ionization, energy pooling, excited-state proton transfer, disproportionation, desorption of preformed ions and thermal ionization. Further secondary ion-molecule reactions then occur such as gas-phase proton-transfer reactions between matrix molecules or between matrix and analyte molecules, gas-phase cationization reactions and electron-transfer reactions[33].

The main purpose of using a matrix in MALDI is to absorb the laser energy. Once the matrix molecules are excited they can use the energy to desorb the analyte molecules. The result is a very soft desorption/ionization technique with little analyte destruction. The matrix also plays an important role as a proton donor or acceptor during

ionization reactions. A low analyte-to-matrix molar ratio (1:1000 to 1:100000) is used to ensure proper desorption and ionization of the analyte.

1.3.2 The Time-of-Flight Mass Analyser The linear time-of-flight (TOF) mass analyser (see Figure 1.5) is the simplest type of mass analyser. It is based on the principle that when an ion with mass m is subjected to a voltage V in the source region, it will be accelerated according to the equation:

$$eV = \tfrac{1}{2} mv^2 \qquad \text{(Equation 1.4)}$$

where v is the ion's final velocity and e is the elementary unit of charge. All ions in the source region will gain the same amount of kinetic energy, so ions with different mass-to-charge ratios will exit the source region with different velocities and separate as they pass through the field-free drift region toward the detector. Since

$$v = \mathrm{D} / t \qquad \text{(Equation 1.5)}$$

where D is the distance travelled and t is the flight time, the different ions will reach the detector at different times. In order to obtain accurate flight times, an exact starting point is necessary. In MALDI-TOF MS the laser pulse used for desorption/ionization is generally used as the starting point. In theory, the masses of ions can now be calculated from their flight times using Equations 1.4 and 1.5. However, since V and D are seldom known with sufficient precision, a calibration curve constructed from the flight times of known mass standards is usually used to calculate the masses of unknowns.

All ions with the same mass-to-charge ratio should reach the detector at the same time, but this is not always so. When the ions form in the plume, they can do so at slightly different times and positions, and with different initial kinetic energies. The temporal and spatial distributions are usually quite small in MALDI, but initial energy distributions can be significant, leading to decreased resolution. Two techniques that can be used to counteract the effect of kinetic energy distributions are time-lag focusing (TLF, also called delayed extraction) and reflectron TOF MS.

A schematic diagram of a time-lag focusing mass spectrometer is shown in Figure 1.6. When the ions are first formed in the source (Figure 1.6a), the repeller plate and the first extraction grid are held at the same voltage, and the newly formed ions are allowed to expand in the MALDI plume. Ions with higher kinetic energies move further away

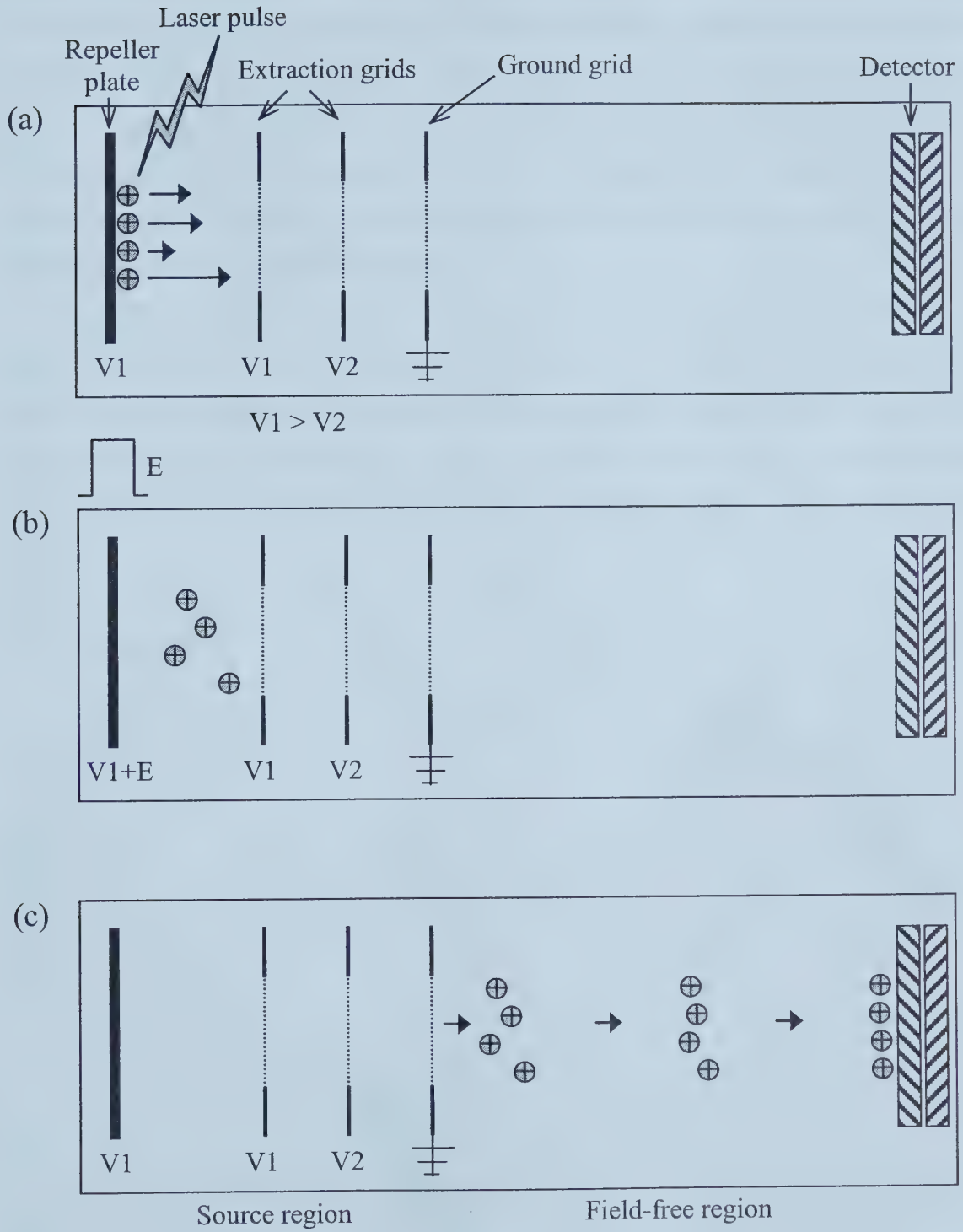

Figure 1.6 Schematic diagram of a time-lag focusing MALDI-TOF mass spectrometer showing a) desorption, b) extraction and c) detection.

from the repeller plate than those with lower energies. After a short delay, an extraction pulse voltage, E, is applied to the repeller plate (Figure 1.6b). The low-energy ions closer to the repeller plate receive more energy from the extraction pulse than the high-energy ions farther away. This allows the low-energy ions to catch up to the high-energy ions as they travel to the detector (Figure 1.6c). Very high resolution can be obtained if the extraction pulse voltage is properly optimized[34]. It has also been demonstrated that if the extraction pulse is modified so that it decreases in amplitude with time, mass accuracy can be improved over a wider mass range[35].

Figure 1.7 shows a MALDI mass spectrometer in reflectron mode, which varies from the linear mode by the addition of a reflectron, or ion mirror. The reflectron is a series of parallel plates to which successively higher potentials are applied. The potential on the last plate (V_R) is usually the same as the source voltage (V_S). Ions with higher initial kinetic energies penetrate deeper into the reflectron field and therefore take a

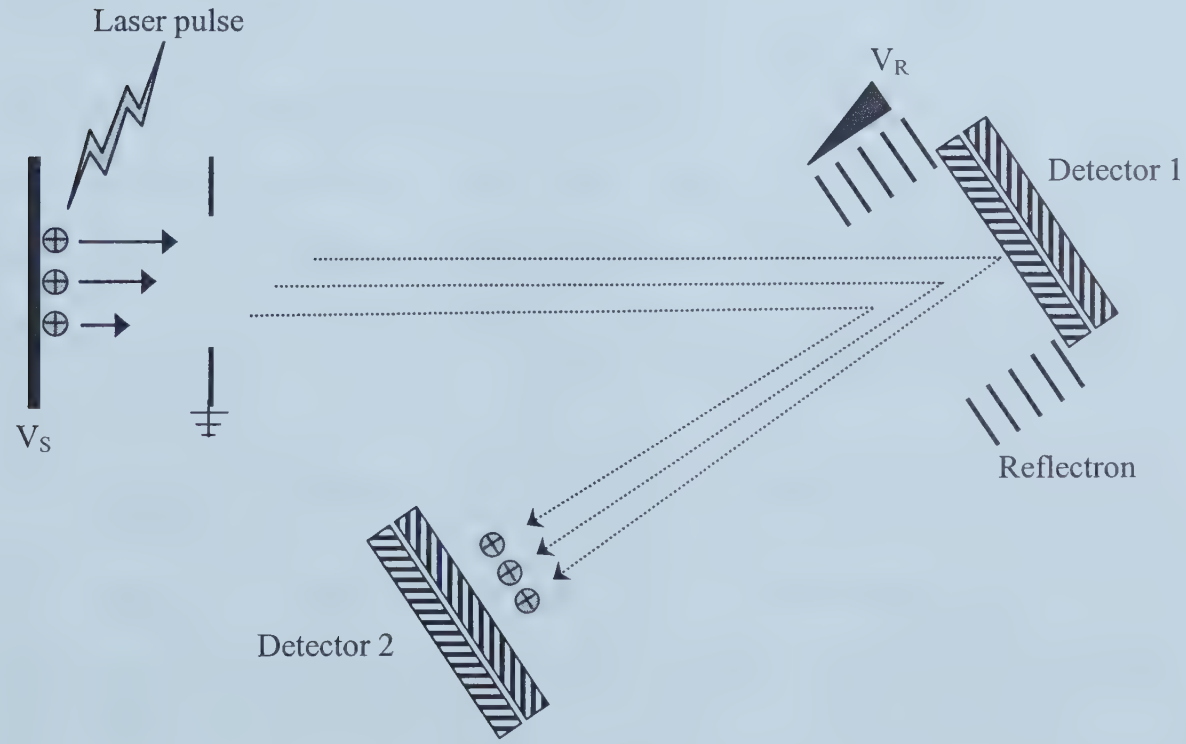

Figure 1.7 Schematic diagram of a MALDI-TOF mass spectrometer in reflectron mode.

longer time to exit the field. As a result, they reach the detector (detector 2) at the same time as the lower-energy ions that have spent less time in the reflectron. While reflectron mode gives better resolution than linear mode, it often has decreased sensitivity because of ion loss during the reflection process. Many MALDI instruments, however, offer the choice of operating in either mode. When desired, the reflectron voltages can be turned off and the instrument can be used in linear mode using detector 1.

The following chapters investigate the use of CE in combination with MALDI MS for proteome analysis. First, the optimization of micro fraction collection for off-line coupling of CE and MALDI MS is discussed. The use of CE/MALDI MS for molecular mass determination and peptide mapping for protein identification is then illustrated. Finally, the application of CE/MALDI MS to the analysis of complex protein mixtures in *Escherichia coli* cell lysates is demonstrated.

1.4 Literature Cited

(1) Roach, J. S. *Anal. Chem.* **2000**, *72*, 609A-611A.

(2) Kheterpal, I.; Mathies, R. A. *Anal. Chem.* **1999**, *71*, 31A-37A.

(3) Anazawa, T.; Takahasshi, S.; Kambara, H. *Anal. Chem.* **1996**, *68*, 2699-2704.

(4) Crabtree, H. J.; Bay, S. J.; Lewis, D. F.; Zhang, J.; Coulson, L. D.; Fitzpatrick, G. A.; Delinger, S. L.; Harrison, D. J.; Dovichi, N. J. *Electrophoresis* **2000**, *21*, 1329-1335.

(5) Larsen, M. R.; Roepstorff, P. *Fres. J. Anal. Chem.* **2000**, *366*, 677-690 and references therein.

(6) Yates, J. R. I. *J. Mass Spectrom.* **1998**, *33*, 1-19 and references therein.

(7) Mann, M.; Hendrickson, R. C.; Pandey, A. *Ann. Rev. Biochem.* **2001**, *70*, 437-473 and references therein.

(8) James, P., Ed. *Proteome Research: Mass Spectrometry*; Springer-Verlag: Berlin, 2001.

(9) Hjerten, S. *Chromatogr. Rev.* **1967**, *9*, 122.

(10) Virtanen, R. *Acta Polytechnica Scand.* **1974**, *123*, 1.

(11) Mikkers, F. E. P.; Everaerts, F. M.; Verheggen, T. P. E. M. *J. Chromatogr.* **1979**, *169*, 11.

(12) Jorgenson, J. W.; Lukacs, K. D. *Anal. Chem.* **1981**, *53*, 1298-1302.

(13) Lee, C. S. In *Handbook of Capillary Electrophoresis*; 2 ed.; Landers, J. P., Ed.; CRC Press: Boca Raton, FL, 1997; pp 717-740.

(14) Heiger, D. N. *High Performance Capillary Electrophoresis- An Introduction*; 3 ed.; Hewlett Packard Company: France, 1997.

(15) Nagra, D. S.; Li, L. *J. Chromatogr. A* **1995**, *711*, 235-245.

(16) Li, L.; Wang, A. P. L.; Coulson, L. D. *Anal. Chem.* **1993**, *65*, 493-495.

(17) Whittal, R. M.; Russon, L. M.; Li, L. *J. Chromatogr. A* **1998**, *794*, 367-375.

(18) Orsnes, H.; Zenobi, R. *Chem. Soc. Rev.* **2001**, *30*, 104-112 and references therein.

(19) Murray, K. K. *Mass Spectrometry Reviews* **1997**, *16*, 283-299 and references therein.

(20) Figeys, D.; Aebersol, R. *Electrophoresis* **1998**, *19*, 885-892.

(21) Strausbauch, M. A.; Wettstein, P. J. In *Handbook of Capillary Electrophoresis*; 2nd ed.; Landers, J. P., Ed.; CRC Press: Boca Raton, 1997; pp 841-864.

(22) Keough, T.; Takigiku, R.; Lacey, M. P.; Purdon, M. *Anal. Chem.* **1992**, *64*, 1594-1600.

(23) Walker, W. L.; Chiu, R. W.; Monnig, C. A.; Wilkins, C. L. *Anal. Chem.* **1995**, *67*, 4197-4204.

(24) Johnson, T.; Bergquist, J.; Ekman, R.; Nordhoff, E.; Schurenberg, M.; Kloppel, K.-D.; Muller, M.; Lehrach, H.; Gobom, J. *Anal. Chem.* **2001**, *73*, 1670-1675.

(25) Preisler, J.; Foret, F.; Karger, B. L. *Anal. Chem.* **1998**, *70*, 5278-5287.

(26) Preisler, J.; Hu, P.; Rejtar, T.; Karger, B. L. *Anal. Chem.* **2000**, *72*, 4785-4795.

(27) Van Veelen, P. A.; Tjaden, U. R.; Van Der Greef, J.; Ingendoh, A.; Hillenkamp, F. *J. Chromatogr.* **1993**, *647*, 367-374.

(28) Zhang, H.; Caprioli, R. M. *J. Mass Spectrom.* **1996**, *31*, 1039-1046.

(29) Tanaka, K.; Ido, Y.; Akita, S. In *Proceedings of the Second Japan-China Joint Symposium on Mass Spectrometry*; Matsuda, H., Liang, X.-T., Eds.; Bando Press: Osaka, Japan, 1987; pp 185-188.

(30) Karas, M.; Bachmann, D.; Bahr, Y.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53-68.

(31) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(32) Karas, M.; Gluckmann, M.; Schafer, J. *J. Mass Spectrom.* **2000**, *35*, 1-12.

(33) Zenobi, R.; Knochenmuss, R. *Mass Spectrometry Reviews* **1998**, *17*, 337-366 and references therein.

(34) Whittal, R. M.; Li, L. *Anal. Chem.* **1995**, *67*, 1950-1954.

(35) Whittal, R. M.; Russon, L. M.; Weinberger, S. R.; Li, L. *Anal. Chem.* **1997**, *69*, 2147-2153.

Chapter 2

Capillary Electrophoresis Using a Surfactant-Treated Capillary Coupled with Offline Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for High Efficiency and Sensitivity Detection of Proteins[a]

2.1 Introduction

One of the main problems when analyzing proteins by CE is analyte adsorption onto the capillary wall. Protein adsorption can lead to serious band broadening, sample loss and poor reproducibility. A common approach to preventing such adsorption is to form a coating inside the capillary[1]. There are two main types of capillary coatings: permanent and dynamic. Permanent coatings are formed by chemically modifying the capillary surface[2], while dynamic coatings are formed through the adsorption of polymers or surfactants[3]. The coating procedures for permanent coatings are usually time-consuming and tedious. Commercial capillaries with permanent coatings are available, but they are expensive and have limited lifetimes and operating conditions. Dynamic coatings, on the other hand, are inexpensive and easy to form by simply rinsing the capillary with a solution containing the surfactant or polymer. More importantly, superior performance has been reported in the prevention of protein adsorption with surfactant-modified capillaries. When coupling CE with MS, however, the use of surfactants has been traditionally avoided due to their incompatibility with MS detection. Surfactants form adducts with proteins and suppress sample ionization[4]. Hence CE/MALDI-MS has generally been performed using permanently coated capillaries.

The use of a semi-permanent capillary wall coating formed by a two-tailed cationic surfactant, didodecyldimethylammonium bromide (DDAB), was recently reported by Melanson *et al.*[5]. Similar to other cationic surfactants, DDAB coating

---

[a] A version of this chapter has been accepted for publication: Ken K.-C. Yeung, Andrea G. Kiceniuk and Liang Li *"Capillary electrophoresis using a surfactant-treated capillary coupled with offline matrix-assisted laser desorption ionization mass spectrometry for high efficiency and sensitivity detection of proteins" J. Chromatogr. A*, in press.

formation takes place by rinsing the capillary with a surfactant solution. However, the two-tailed nature of DDAB results in a very strong hydrophobic effect so that the coating remains after flushing out the DDAB solution. The coated capillary is then filled with a surfactant-free buffer in which electrophoresis takes place. This potentially eliminates the presence of surfactant in the bulk solution inside the capillary, and therefore the surfactant should not interfere with the MALDI-MS measurement. The first application of surfactant-modified capillaries in CE/MALDI-MS is demonstrated and studied in this chapter.

As discussed in Chapter 1, the proteins separated by CE are collected in fractions before further manipulation and analysis. Electrokinetic fraction collection is preferred because it allows a high separation efficiency to be maintained. Electrokinetic fractionation is not as straightforward as hydrodynamic fractionation, which requires only the application of pressure at the capillary inlet. During electrokinetic collection, an electrolyte is required in the collection vial to maintain the electrical current between the electrode and the capillary outlet. The electrophoretic buffer is generally used as the electrolyte. However, metal cations from the buffer, such as sodium and potassium, can drastically suppress the MALDI-MS signal of proteins. This suppression is particularly problematic when analyzing low concentration proteins or peptides[6]. For this reason, fraction collection for CE/MALDI-MS is usually performed hydrodynamically[7,8]. The use of dilute acid instead of electrophoretic buffer as the fraction collection electrolyte has been reported to improve MALDI signals[9]. This chapter illustrates that CE fractionation with a dilute acid can be performed electrokinetically in an easy and reliable manner. This technique allows multiple consecutive fraction collections in one CE separation.

2.2 Experimental

2.2.1 Materials. Nanopure 18MΩ water (Barnstead) was used in the preparation of all solutions. Reagent grade orthophosphoric acid (Mallinckrodt, Paris, KY), acetic acid (Anachemia, Montreal, PQ, Canada), or formic acid (Fisher Scientific, Fair Lawn, NJ) was used to prepare the CE buffers. The pH was adjusted with reagent grade sodium

hydroxide (Fisher). Hydrochloric acid used in fraction collection was purchased from Anachemia. Didodecyldimethylammonium bromide (DDAB) and the standard proteins cytochrome *c* (bovine heart, C3131), cytochrome *c* (horse heart, C7752), and lysozyme (chicken egg white, L6876) were used as received from Sigma-Aldrich Canada (Oakville, ON). α-Cyano-4-hydroxycinnamic acid (HCCA, Sigma-Aldrich) was purified by recrystallization from ethanol prior to use.

2.2.2 CE Separation for Proteins. An Agilent ^{3D}CE instrument (Palo Alto, CA) equipped with an ultraviolet-visible diode array detector was used to perform all CE separations and fraction collections. Fused silica capillaries (Polymicro Technologies, Phoenix, AZ) with an inner diameter of 50 μm, an outer diameter of 365 μm, an effective length of 22.3 cm and a total length of 30.5 cm were used. The capillary was thermostated at 25°C in all experiments. ChemStation software (Rev. A.06.03, Agilent) was used for data acquisition and instrument control on a Pentium based microcomputer. Detection of all proteins was performed by direct UV absorption at 192 nm. Hydrodynamic sample injection was performed at a constant pressure of 5 kPa.

Capillary pre-conditioning was performed with a high-pressure (100 kPa) rinse with NaOH for 5 min followed by distilled water for 1 min. New capillaries were used for each buffer system to avoid cross-contamination. Capillary coating with DDAB was performed with a 5-min rinse (100 kPa) with 0.1 mM DDAB in water[5]. A 1-min rinse (100 kPa) with the electrophoretic buffer was used to remove the excess surfactant. Between subsequent separations, the capillary was reconditioned (100 kPa) for 2 minutes with 0.1 mM DDAB and for 1 minute with the electrophoretic buffer.

For the fraction collection solution study, samples of 0.2 μg/μL horse heart cytochrome *c* in water were injected hydrodynamically at 5 kPa for 2.0 sec in all cases. Phosphate, acetate, or formate, 25 mM at pH 4.00, were used as the electrophoretic buffers, and the applied voltage was –6 kV. In the detection limit study, 0.2 μg/μL horse heart cytochrome *c* was injected at 5 kPa for 1.0 sec. A phosphate buffer, 25 mM at pH 4.00, was used and the applied voltage was –6 kV.

In the separation and fraction collection of standard protein mixtures, a mixture of bovine heart cytochrome *c*, horse heart cytochrome *c*, and lysozyme (0.2 μg/μL each)

was injected at 5 kPa for 2.0 sec. The applied voltage in this case was –12 kV, and a 10 mM phosphate buffer at pH 7.00 was used. Efficiencies of all CE peaks were determined using the tangent method.

2.2.3 CE Fraction Collection. Collection of analyte bands at the outlet of the capillary was performed electrokinetically into polypropylene micro inserts (Agilent, Part no. 5182-0549). In the fractionation solution study, various buffers and dilute acid were placed in the insert for electrokinetic fraction collection. To minimize dilution, the volume of the fractionation solution was limited to 5 μL or less. This 5-μL volume resulted in a solution depth of *ca.* 2 mm in the micro insert. To provide electrical contact between the outlet electrode and the fractionation solution, the top of the micro insert was trimmed by 5 mm in order to raise the position of the insert closer to the electrode and capillary.

The automatic fraction collection function available in the ChemStation software (Agilent) was not used in the experiment because it did not account for the vial exchange time (6 sec) during fraction collection. This led to erroneous fraction collection timing, particularly during short fractionations (< 30 sec) for the highly efficient protein peaks. As a result, the fraction collection was performed manually.

The timing of fraction collection is calculated assuming that the mobility of the analyte is constant during the separation. That is, the time at which an analyte band reaches the outlet of the capillary (t_{outlet}) is obtained by multiplying the recorded migration time (t_m) with the ratio of total capillary length (L_t) to the length from inlet to detector (L_d):

$$t_{outlet} = t_m \cdot \frac{L_t}{L_d} \qquad \text{(Equation 2.1)}$$

At t_{outlet}, the CE separation is interrupted by switching off the applied field. The electrophoretic buffer vial at the capillary outlet is removed, and the fractionation micro insert is placed at the outlet. The applied field is then resumed to perform electrokinetic fraction collection. The time length of electrokinetic fraction collection is determined by the baseline width of the peak at the capillary outlet ($W_{b,outlet}$), which is calculated from the baseline width recorded at the detector ($W_{b,detector}$):

$$W_{b,outlet} = W_{b,detector} \cdot \frac{L_t}{L_d} \qquad \text{(Equation 2.2)}$$

When needed, consecutive fractions can be collected by placing new micro inserts for fractionation. Otherwise, the electrophoretic buffer vial can be repositioned at the outlet to resume the CE separation.

2.2.4 Sample Deposition and MALDI Measurement for Protein Standards. Mass spectral measurements were conducted with a time-lag focusing (TLF) MALDI-TOF mass spectrometer constructed in-house[10,11]. Mass spectral data acquisition and processing was performed with Hewlett Packard supporting software HP DADA-601 (S.02.99.1101). Data was reprocessed using IGOR Pro (Version 3.13, Wavemetrics Inc., Lake Oswego, OR).

A three-layer sample deposition method was used for MALDI protein analysis. This method generated reproducible mass spectra from a given sample. In this method, 1 μL of 10 mg/mL HCCA in 20% methanol/acetone v/v was first deposited on the MALDI target and allowed to dry to form a thin layer of fine crystals. Then 0.25 μL of a saturated HCCA solution in 33% methanol/water v/v was deposited on top of the first layer and allowed to dry. Finally, a 0.25 μL aliquot of the sample was spotted as a third layer and allowed to dry prior to measurement. Each analysis was repeated in triplicate, with an average of 75 laser shots summed for each mass spectrum.

2.3 Results and Discussion

2.3.1 Selection of CE Buffer and Fraction Collection Solution. Three CE buffer systems were evaluated for use with the DDAB-coated capillaries: 25 mM phosphate, formate and acetate, all at pH 4.00. Cytochrome *c* (horse heart) was used as a test protein, with *ca.* 92 fmol injected for each run. The electropherograms obtained with the different buffers are shown in Figures 2.1A-C respectively. The peak efficiencies are 600,000 plates/m in phosphate (Figure 2.1A), 300,000 plates/m in formate (Figure 2.1B) and 200,000 plates/m in acetate (Figure 2.1C). The measured efficiencies are similar to those reported previously by Melanson *et al.*[5].

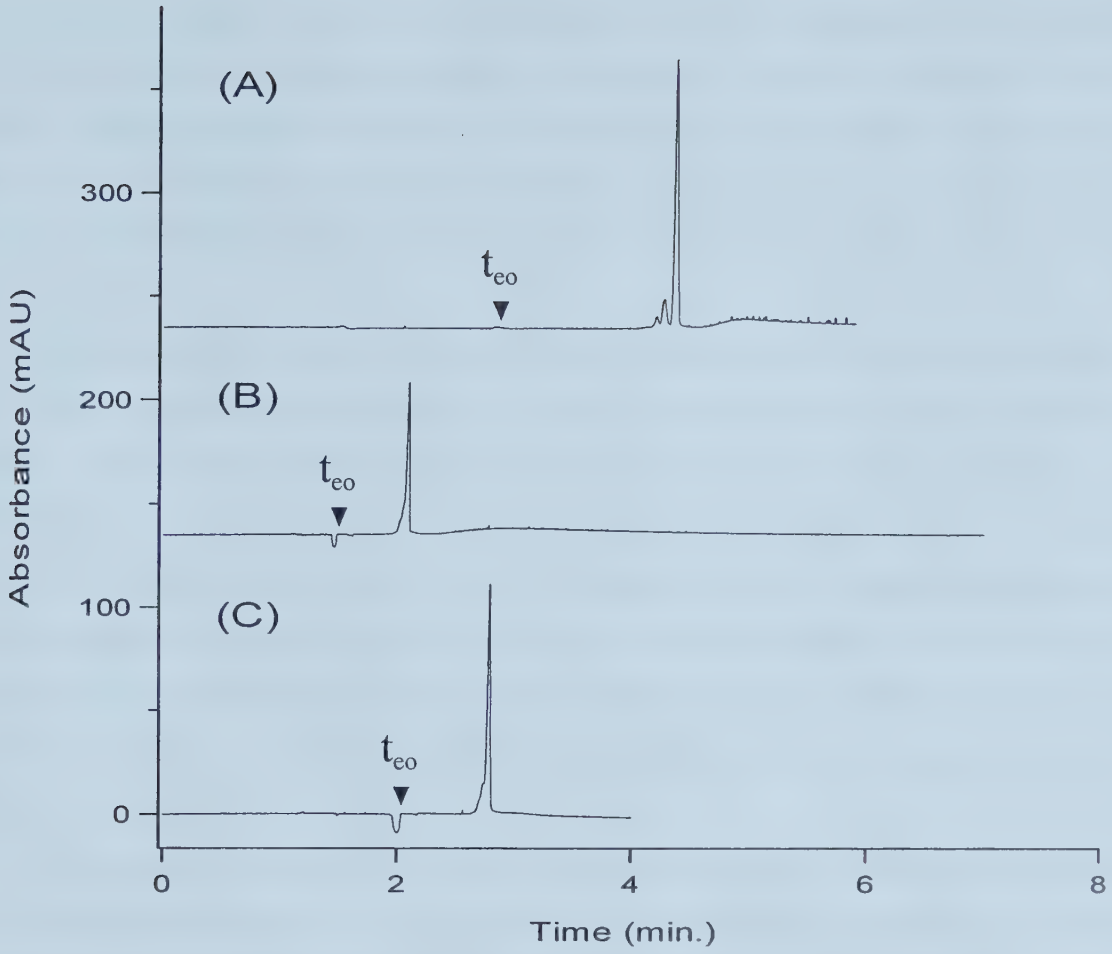

Figure 2.1 CE electropherograms of horse heart cytochrome *c* in (A) sodium phosphate, (B) sodium formate, and (C) sodium acetate buffers (25 mM, pH 4.0). Separation is performed at -6 kV in 50 μm I.D., 22.3 cm (effective)/30.5 cm (total) DDAB-coated capillaries. The EOF is determined by injection of acetonitrile (A) or water (B-C).

The migration times of cytochrome *c* in all three buffers are relatively short, from 2.0 minutes in formate to 4.4 minutes in phosphate. The short migration times are caused by the fast electroosmotic flow (EOF) generated in the DDAB-modified capillary. Most of the available coatings, either dynamic or permanent, are neutral in charge. The EOF generated with these coatings is therefore either greatly suppressed or near zero. The mobilities of analytes are solely driven by their own charges in the absence of the EOF. Slightly charged or near-neutral proteins therefore migrate very slowly, resulting in long analysis times. DDAB, on the other hand, is a cationic surfactant that generates a highly cationic capillary wall coating, and thus a fast anodic EOF. Although cationic proteins

migrate in the opposite direction as the anodic EOF, the magnitude of the EOF is sufficient to sweep the counter-migrating proteins to the detector. Such fast EOF is useful in carrying proteins with low or zero mobility to the detector, enabling their detection that would not otherwise be possible. The variation in the migration time among different buffers (Figure 2.1) is due to the different magnitude of EOF in the various buffers[12,13].

Micro fraction collection was performed as outlined in the Experimental section. The measured peak width at baseline is *ca.* 6.5 seconds for cytochrome *c* in all three buffers. Taking into account the detector-to-total length ratio, this translates to an electrokinetic fraction collection interval of 9.0 seconds based on Equation 2.2. For each of the three buffer systems (Figures 2.1A-C), fraction collection was performed first in 5 μL of the electrophoretic buffer and then in 5 μL of dilute (10 ppm) HCl solution. The purpose of using a dilute HCl solution is to maintain an almost sodium-free electrolyte condition that will not adversely affect the performance of the subsequent MALDI experiment (see below).

The mass spectra of the cytochrome *c* fractions collected in electrophoretic buffer are presented in Figures 2.2A, 2.3A and 2.4A for phosphate, formate, and acetate respectively. The corresponding mass spectra of the fractions collected in HCl are shown in Figures 2.2B, 2.3B and 2.4B. As a control experiment, the mass spectrum of a 20 nM cytochrome *c* solution in water was also recorded. The 20 nM concentration was selected to approximate the 92 fmol of protein collected in 5 μL. The mass spectra of this solution represent the optimal signal intensity of cytochrome *c* in the absence of any signal suppression or sample loss during fraction collection. Since the MALDI measurements for the three buffers were performed on three different days, a control measurement of 20 nM cytochrome *c* was performed on each day. They are shown in Figures 2.2C, 2.3C and 2.4C. All spectra shown in Figures 2.2-2.4 were collected in triplicate. The spectra shown represent the average of the triplicates.

In all of the spectra recorded (Figures 2.2-2.4), three main ion peaks are detected: $(M+H)^+$, $(M+2H)^{2+}$, and $(M+3H)^{3+}$. A variation in signal intensity is observed between the different fractionation electrolytes. For all three of the buffers studied, the mass spectra of fractions collected in the electrophoretic buffer result in the lowest signal

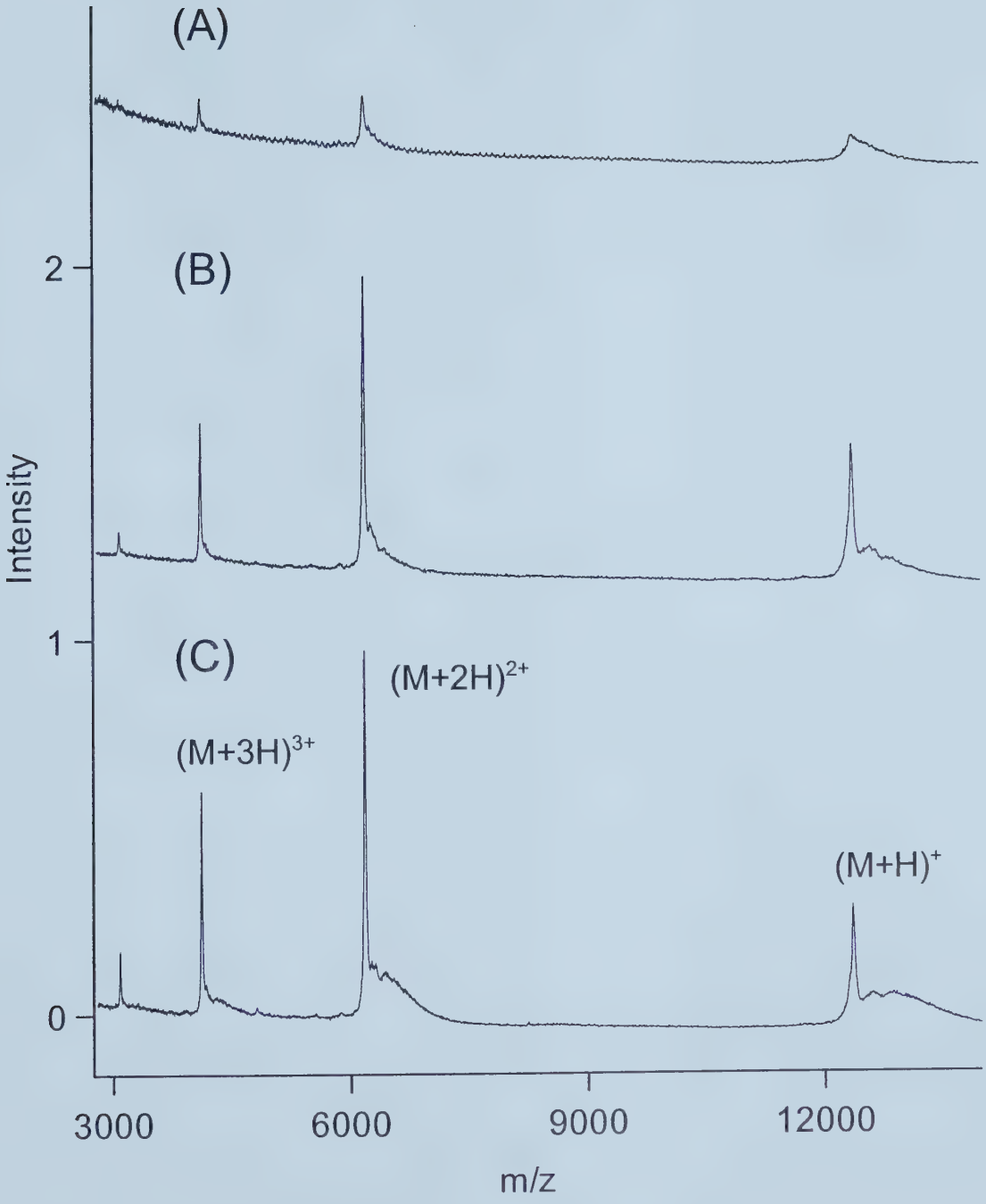

Figure 2.2 MALDI mass spectra of cytochrome *c* from (A) CE fraction collected in electrophoretic buffer, (B) CE fraction collected in 10 ppm HCl, and (C) a 20 nM standard solution. CE separation was performed in 25 mM sodium phosphate at pH 4.0.

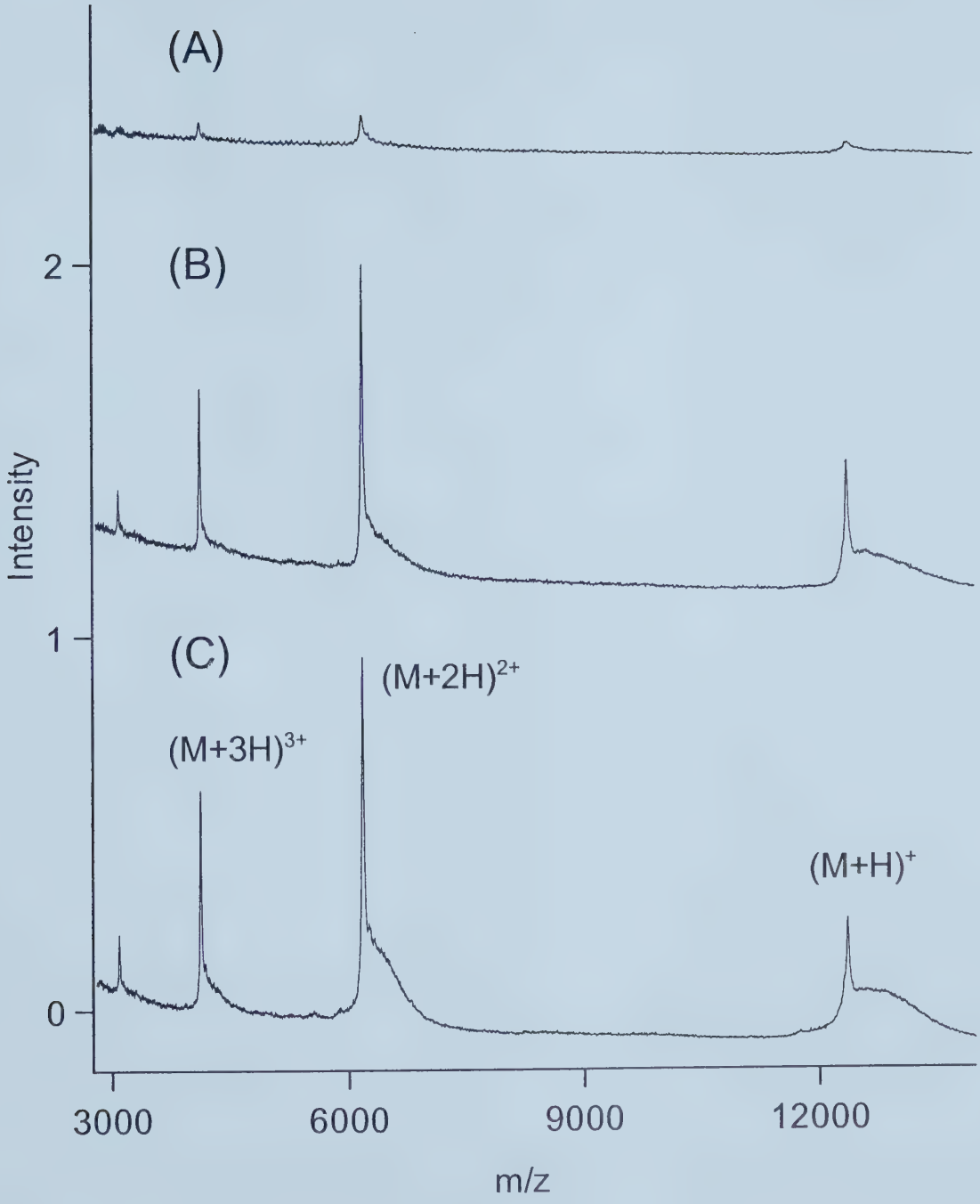

Figure 2.3 MALDI mass spectra of cytochrome *c* from (A) CE fraction collected in electrophoretic buffer, (B) CE fraction collected in 10 ppm HCl, and (C) a 20 nM standard solution. CE separation was performed in 25 mM sodium formate at pH 4.0.

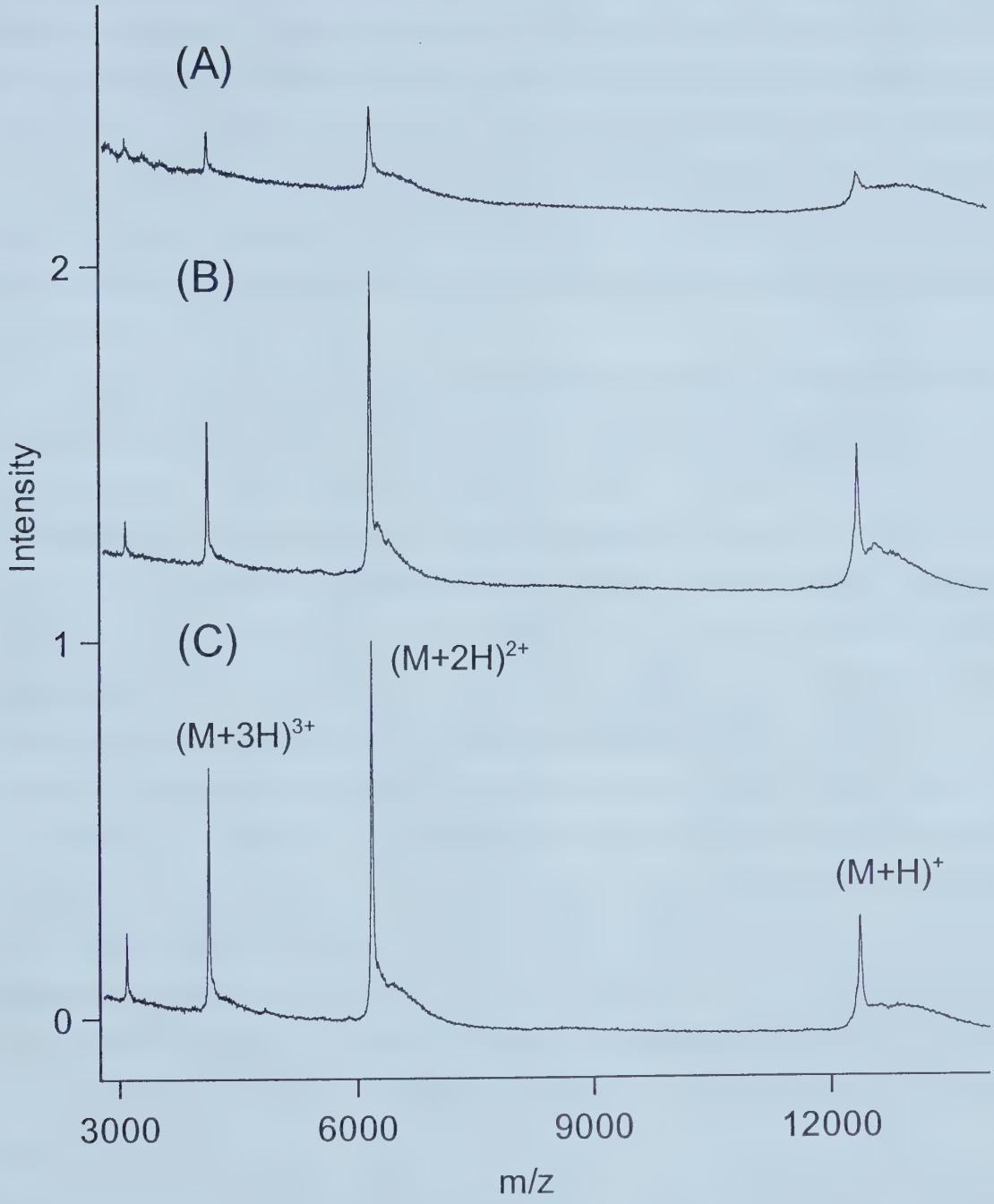

Figure 2.4 MALDI mass spectra of cytochrome *c* from (A) CE fraction collected in electrophoretic buffer, (B) CE fraction collected in 10 ppm HCl, and (C) a 20 nM standard solution. CE separation was performed in 25 mM sodium acetate at pH 4.0.

intensity (Figures 2.2A, 2.3A and 2.4A). When the same fraction collection is performed in dilute HCl, peak intensities similar to those from the 20 nM cytochrome *c* standard solution are observed. This confirms that MS signal suppression occurs when the electrophoretic buffers (sodium phosphate, sodium formate and sodium acetate) are used as the fraction collection electrolyte. As a result, the dilute HCl solution was used for fraction collection in all subsequent experiments to minimize MS signal suppression. Since the choice of running buffer was no longer a concern for fraction collection, phosphate buffer, which gave the highest separation efficiency, was used for subsequent experiments.

A potential problem in using a non-buffered electrolyte, such as dilute HCl, for electrokinetic fraction collection is electrolysis. Electrolysis at the capillary outlet can cause a pH change, which in turn may alter the protein mobility. The use of a DDAB modified capillary helps to reduce the effect of the potential electrolysis. A fast, forward-moving EOF is generated in DDAB capillaries, which may help prevent the fraction collection electrolyte from entering the capillary at the outlet. In addition, a reversed polarity setting is used with a DDAB capillary, *i.e.* the anode is at the capillary outlet. If electrolysis takes place, the pH of the fraction collection electrolyte will be lowered by oxidation. CE separation of proteins is normally performed in acidic buffers, such as the pH 4 buffer used in this work. Fortunately, the ionization of proteins is less sensitive to pH changes in the acidic direction, as these proteins are close to being fully protonated at pH 4. On the other hand, normal polarity must be used for most permanently derivatized capillaries due to the suppressed EOF. Reduction takes place at the capillary outlet and increases the buffer pH. This can cause significant changes in the proteins' ionization and thus mobilities, which in turn affects the quality of separation and the accuracy of fraction collection.

2.3.2 Compatibility of DDAB with MALDI-MS. Although DDAB is useful for achieving fast analysis times and allows the use of non-buffered fraction collection electrolytes, it is important to make sure that the DDAB from the capillary coating does not cause interference in MALDI-MS. The MALDI spectra in panel B of Figures 2.2-2.4 are recorded from fractions collected from DDAB-coated capillaries. The intensities of the $(M+2H)^{2+}$ peaks are 80-90% of those observed in the spectra of 20 nM cytochrome *c*

solution (panel C in Figures 2.2-2.4). The small difference in intensity may reflect the different amount of cytochrome *c* used, 92 fmol from the CE fraction and 100 fmol from the 20 nM solution. In addition, sample handling during fraction collection, such as solution transfer between vials, can also result in sample loss. Overall, significant interference caused by DDAB is not evident. The first coupling of CE and MALDI-MS with the use of a surfactant modified capillary is therefore successfully demonstrated. The use of permanently derivatized capillaries is no longer the only choice for CE/MALDI-MS.

2.3.3 CE/MALDI-MS Detection Limit. To determine the detection limit of our MALDI-MS measurements on the CE fractions collected in 10 ppm HCl, CE/MALDI-MS was performed with different amounts of cytochrome *c* (horse heart) injected. The separation was performed in phosphate buffer (25 mM) at pH 4.00, and fraction collection was performed as previously described at –6 kV for 9.0 seconds. The signal-to-noise ratio (S/N) of the $(M+2H)^{2+}$ peak was recorded for each sample in four replicates. The lowest amount of cytochrome *c* (horse heart) injected that resulted in an MS signal of S/N 3 in all four replicates is 23 fmol, with S/N ranging from 6 to 60. The wide range of signals observed from the replicate experiments indicates that this amount is close to the detection limit. Indeed, any lower amount of cytochrome *c* resulted in signals below the detection limit in at least one of the four replicates. The lowest detection limit reported in the literature for CE/MALDI-MS is 2 fmol for proteins injected in CE[14]. Our detection limit is about 10 times higher. Unlike the direct sample deposition used in reference 14, not all of the protein collected in each fraction is used for MALDI-MS in our case. Out of the 5 μL in each fraction, only 0.25 μL and thus only 1.2 fmol is actually spotted onto the MALDI probe for each measurement. The detection limit of our laboratory-built MALDI instrument is around 1 fmol for proteins such as cytochrome *c*[15]. In other words, the MALDI-MS detection limit is not affected by the CE separation and fraction collection.

Since only 0.25 μL is consumed for each MALDI-MS measurement, one can, in principle, perform 20 measurements for each 5-μL CE fraction. Alternatively, the 5-μL fraction can be divided into several portions for various sample treatments such as

enzymatic digestions and chemical reactions. Chapter 3 discusses the use of tryptic digestion for protein identification and structure analysis.

The amount of fractionation electrolyte can be reduced to less than 5 μL in order to lower the detection limit. It was found that the smallest fraction volume that could be used was 2 μL. In theory this can lower the detection limit from 23 fmol to 9 fmol, which is closer to the low detection limit record of 2 fmol reported in the literature[14]. Nevertheless, the 2-μL electrolyte volume is so small that it occasionally does not result in a successful collection, presumably when it fails to maintain electrical contact between the electrode and the capillary outlet. Care must be exercised when using such small volumes of fractionation electrolyte.

Concentrating the collected CE fraction prior to MALDI measurement was also investigated as a method for increasing sensitivity. A 5-μL fraction was transferred to a siliconized polypropylene vial (600 μL) and evaporated down to 1 μL by vacuum centrifugation. Unfortunately, this attempt at concentration was not successful as the protein was lost, presumably due to adsorption onto the container. As a result, the 'concentrated' samples actually resulted in lower MALDI signals. In the future, other concentration techniques, such as the use of reverse phase particles[16], could be investigated to increase sensitivity further.

2.3.4 Accuracy of Fraction Collection Timing. Since the separation efficiency in this CE method is very high, a suitable means of collecting fractions for only a few seconds is required. The timing of our electrophoretic fraction collection is based on the extrapolation of the migration time recorded by an online detector. Errors may arise from any changes in analyte mobility during electrophoresis and fraction collection. Electrolysis can take place at the electrode and alter the pH, protein mobility and EOF. Additional band broadening or siphoning during vial transfers, as well as the uncertainty of the voltage control during ramping, can also give rise to erroneous timing during fraction collection.

To evaluate the accuracy of fraction collection, multiple consecutive fraction collections were performed within one CE separation. The CE separation of a mixture of three proteins, cytochrome *c* (bovine heart), cytochrome *c* (horse heart) and lysozyme

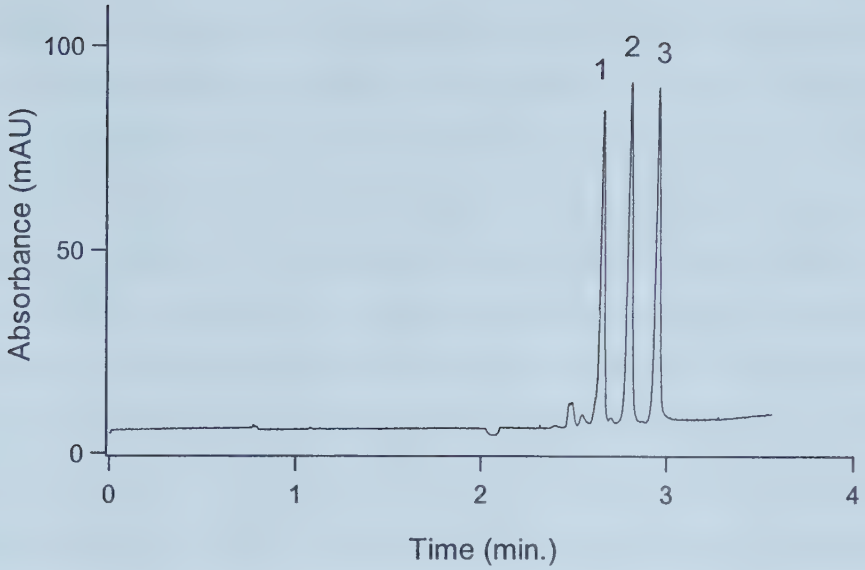

Figure 2.5 CE separation of a mixture of bovine heart cytochrome *c* (1), horse heart cytochrome *c* (2), and lysozyme (3). Separation is performed at -12 kV in a 50 μm I.D., 22.3 cm (effective)/30.5 cm (total) DDAB-coated capillary with 10 mM sodium phosphate buffer, pH 7.0.

(chicken egg white), was performed at –12 kV in a pH 7.0 phosphate buffer (10 mM). The proteins and the experimental conditions were optimized to yield three closely resolved peaks (Figure 2.5). Poor accuracy in the fractionation timing would result in the collection of unintended components, which would be reflected in the subsequent MALDI mass spectra. Based on the calculation from Equations 2.1 and 2.2, collection of the three protein peaks was performed as three consecutive fractions with intervals of 10-12 seconds at -12 kV. MALDI mass spectra were recorded for the fractions. Initially the MALDI signals appeared to be weaker than that recorded previously when the fractionation voltage was -6 kV. A similar observation has been reported by others[17]. It is caused by a loss of analytes when high voltage (-12 kV) is used in electrokinetic fraction collection. Sample binding or interaction with the outlet electrode occurs, resulting in incomplete collection. The problem can be solved by simply reducing the fractionation voltage and adjusting the collection time duration proportionally[18]. Therefore, fractionation was performed at half the voltage (- 6 kV) for twice the time duration (18 seconds). The intensity of the MS signal improved approximately twofold with the reduced fractionation voltage.

The MALDI mass spectra of the cytochrome *c* (bovine heart), cytochrome *c* (horse heart), and lysozyme CE fractions are presented in panels A-C respectively of Figure 2.6. In general, the major peaks observed in each mass spectrum correspond to the protein expected. In the first fraction (Figure 2.6A), only signals from cytochrome *c* (bovine heart) are observed. In the second fraction (Figure 2.6B), small quantities of lysozyme, $(M+2H)^{2+}$, is observed. A small amount of cytochrome *c* (bovine heart) may also be present in the second fraction, but the MS resolution is not sufficient to resolve the small mass difference between the two types of cytochrome *c*. Likewise, small signals from cytochrome *c* (bovine heart) and cytochrome *c* (horse heart) are detected in the third fraction (Figure 2.6C). A small degree of sample carry-over from one fraction to another is suspected to be the cause of the contamination.

The above results show that multiple consecutive fraction collections can be performed within a short collection time and collection of the intended component is evident in the MALDI-MS results. In fact, multiple consecutive fractionations can be performed throughout the entire separation, as illustrated in Chapter 4 with the analysis of *Escherichia coli* cell lysates. Using dilute HCl as fraction collection electrolyte did not lead to any significant error in the collection timing, indicating that the mobility of the proteins was not altered significantly by electrolysis during fractionation. Overall, fraction collection was performed in a precisely controlled manner.

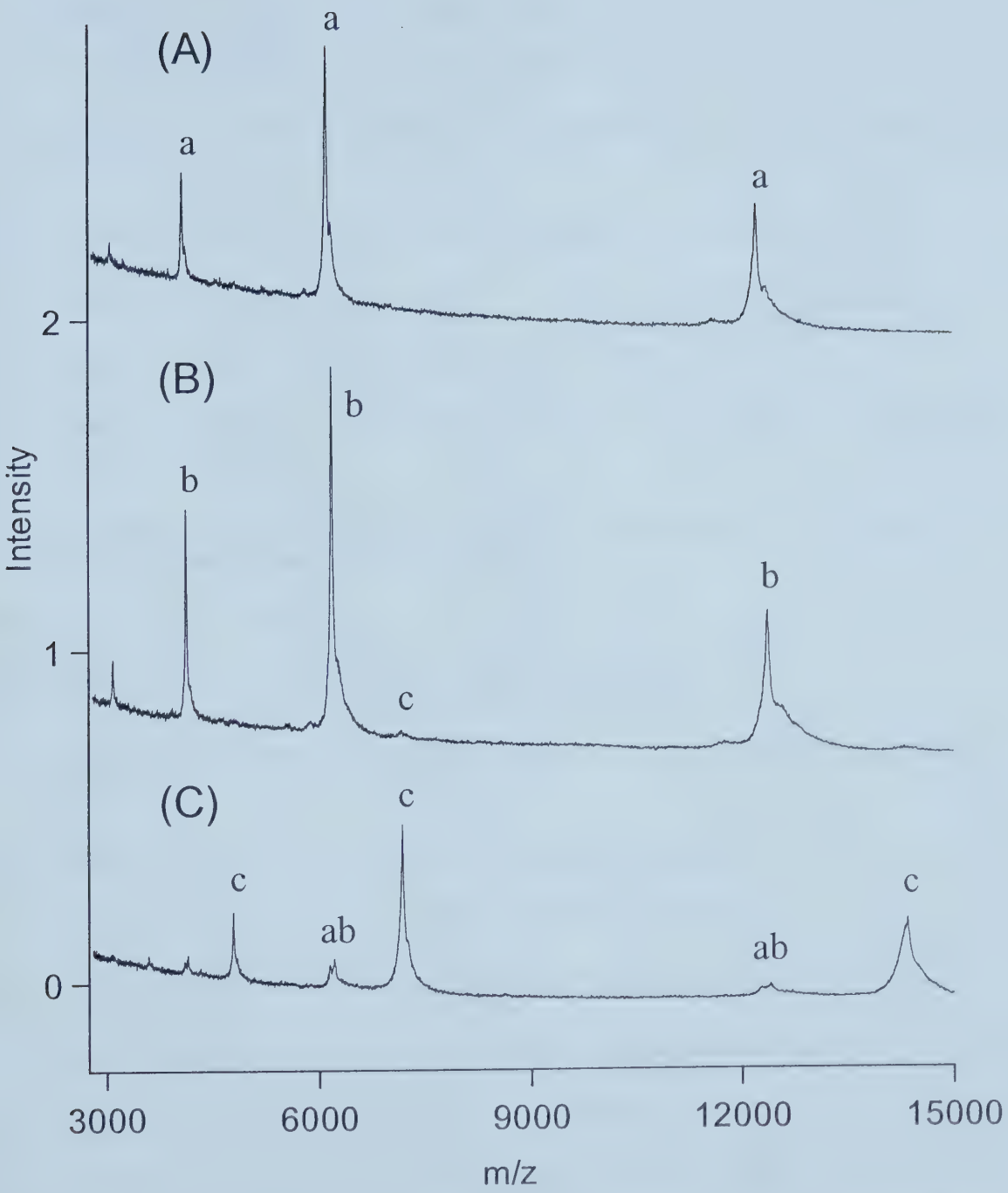

Figure 2.6 MALDI mass spectra of CE fractions for peaks 1, 2 and 3 from Figure 2.5 (A-C, respectively). The lower case letters refer to the peaks identified for bovine heart cytochrome *c* (a), horse heart cytochrome *c* (b), and lysozyme (c).

2.4 Literature Cited

(1) Rodriguez, I.; Li, S. F. Y. *Anal. Chim. Acta* **1999**, *383*, 1-26.

(2) Horvath, J.; Dolnik, V. *Electrophoresis* **2001**, *22*, 644-655.

(3) Righetti, P. G.; Gelfi, C.; Verzola, B.; Castelletti, L. *Electrophoresis* **2001**, *22*, 603-611.

(4) Zhang, N.; Doucette, A.; Li, L. *Anal. Chem.* **2001**, *73*, 2968-2975.

(5) Melanson, J. E.; Baryla, N. E.; Lucy, C. A. *Anal. Chem.* **2000**, *72*, 4110-4114.

(6) Keller, B. O.; Li, L. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 88-93.

(7) Chakel, J. A.; Erno Pungor, J.; Hancock, W. S.; Swedberg, S. A. *J. Chromatogr. B* **1997**, *689*, 215-220.

(8) Choudhary, G.; Chakel, J.; Hancock, W.; Torres-Duarte, A.; McMahon, G.; Wainer, I. *Anal. Chem.* **1999**, *71*, 855-859.

(9) Bergman, A. C.; Bergman, G. *J. Prot. Chem.* **1997**, *16*, 421-423.

(10) Whittal, R. M.; Li, L. *Anal. Chem.* **1995**, *67*, 1950-1954.

(11) Whittal, R. M.; Russon, L. M.; Weinberger, S. R.; Li, L. *Anal. Chem.* **1997**, *69*, 2147-2153.

(12) Lucy, C. A.; Underhill, R. S. *Anal. Chem.* **1996**, *68*, 300-305.

(13) Baryla, N. E.; Lucy, C. A. *Anal. Chem.* **2000**, *72*, 2280-2284.

(14) Zhang, H.; Caprioli, R. M. *J. Mass Spectrom.* **1996**, *31*, 1039-1046.

(15) Keller, B. O.; Ph.D. thesis; The University of Alberta: Edmonton, 2001.

(16) Doucette, A.; Craft, D.; Li, L. *Anal. Chem.* **2000**, *72*, 3355-3362.

(17) Lee, H. G.; Desiderio, D. M. *J. Chromatogr.* **1994**, *686*, 309-317.

(18) Strausbauch, M. A.; Wettstein, P. J. In *Handbook of Capillary Electrophoresis*; 2nd ed.; Landers, J. P., Ed.; CRC Press: Boca Raton, 1997; pp 841-864.

Chapter 3

Digestion Techniques for Peptide Mapping of Capillary Electrophoresis-Separated Proteins

3.1 Introduction

In Chapter 2, CE-separated proteins are identified by MALDI MS based on their molecular masses. Standard proteins are used, so it is easy to differentiate between them based on their known masses. When dealing with unknown proteins, however, molecular mass alone is not enough to unambiguously identify an analyte.[1] As mentioned in Chapter 1, peptide mapping is one method that can be used to identify an unknown protein. The protein is digested, the resulting peptide mixture is analysed by MS and the peptide masses are compared against theoretical digests of proteins until a matching peptide profile is found. Since it is the peptides that are used to identify the protein, the success of peptide mapping depends on an efficient digestion that yields peptides covering as much of the protein's sequence as possible.

At high substrate concentrations, successful digestions are easy to achieve. Many peptides are generated in high concentration, giving strong peptide signals in MS analysis. Problems such as enzyme autolysis, sample loss due to adsorption and signal suppression from contaminants rarely arise. As the substrate concentration is lowered, however, performing digestions becomes more difficult. The reaction becomes first-order with respect to substrate concentration[2], so the lower the substrate concentration, the slower the reaction and the fewer peptides produced. As a result, less-abundant peptides can fall below MS detection limits and sequence coverage often decreases. Furthermore, an increased enzyme/protein ratio is often needed. Under such conditions, autolysis can become more prevalent than substrate digestion. Reaction conditions must be carefully optimized to achieve the best digestion results possible.

This chapter explores three digestion techniques applied to CE-separated proteins at low-femtomolar concentrations. Two of the methods, in-capillary and on-target

digestion, incorporate sample concentration steps prior to digestion to help decrease the problems caused by low substrate concentrations. These methods are compared to the commonly used bulk solution digestion method. The relative merits of each method are discussed, as well as some fundamental obstacles that arise.

3.2 Experimental

3.2.1 Materials. All aqueous solutions were prepared in Nanopure 18MΩ water (Barnstead). Buffers were prepared from reagent grade orthophosphoric acid (Mallinckrodt, Paris, KY), and the pH was adjusted with reagent grade sodium hydroxide (Fisher Scientific, Fair Lawn, NJ). Hydrochloric acid used in fraction collection was purchased from Anachemia (Montreal, PQ). Didodecyldimethylammonium bromide (DDAB), α-cyano-4-hydroxycinnamic acid (HCCA), spectrophotometric grade trifluoroacetic acid (TFA) and the standard proteins trypsin (bovine pancreas), cytochrome *c* (horse heart) and bovine serum albumin (BSA) were purchased from Sigma-Aldrich Canada (Oakville, ON). Reagent grade acetonitrile and methanol were purchased from Fisher. All reagents were used as received except for the HCCA, which was purified by recrystallization from ethanol prior to use.

3.2.2 CZE Protein Separation and Fraction Collection. All CZE separations and fraction collections were performed with an Agilent ^{3D}CE instrument (Palo Alto, CA) equipped with an ultraviolet-visible diode array detector. Fused silica capillaries (Polymicro Technologies, Phoenix, AZ) with an inner diameter of 50 μm, an outer diameter of 365 μm, and a total length of 30.5 cm (22.3 cm to the detector) were used. Capillaries were coated with DDAB prior to use and reconditioned between subsequent separations as described in Chapter 2. A pH 4.0 sodium phosphate buffer (25 mM) was used in all experiments. The applied voltage used during electrophoresis was –6 kV, and the capillary was thermostated at 25°C. Sample injection was performed hydrodynamically at 50 mbar (5 kPa) for 1.0 or 2.0 seconds, depending on the sample concentration and desired loading. Data acquisition and instrument control was

performed using ChemStation software (Rev. A.06.03, Agilent) on a Pentium-based microcomputer. Direct UV absorbance detection was performed at 192 nm.

Fraction collection was performed electrokinetically into polypropylene micro inserts (Agilent, Part no. 5182-0549). As described in Chapter 2, the fraction collection timing was calculated based on the analyte's migration time and the ratio between the total and effective lengths of the capillary (Equation 2.1). A 3-μL volume of dilute (10 ppm) HCl was placed in the micro insert to maintain electrical contact between the capillary and electrode during fraction collection. The top 5 mm of the micro-insert was trimmed off to raise the position of the insert and thus bring the electrode and capillary outlet closer to the bottom of the micro-insert.

3.2.3 Bulk Solution Digests. Bulk digests were carried out in 0.6 mL siliconized polypropylene vials (Rose Scientific, Edmonton, AB). For each digest, 5 μL of the desired concentration of cytochrome *c* solution was mixed with 1 μL of trypsin solution and 5 μL of 50 mM ammonium bicarbonate (to maintain a suitable pH for digestion). The vials were incubated at 37°C for 30 minutes, then the digestion was quenched with ~0.1 μL of 1% TFA. The digested samples were stored on ice and spotted for MALDI analysis as soon as possible.

A two-layer sample deposition procedure[3] was used for MALDI analysis of the digests. The MALDI target was prepared by depositing 1 μL of the first-layer matrix solution (10 mg/mL HCCA in 20% methanol/acetone v/v) and allowing it to dry and form a thin layer of fine crystals. The sample was then mixed 1:1 with the second-layer matrix solution (33% methanol/0.1% TFA/water v/v/v saturated with HCCA), and 0.25 μL of the mixture was deposited on top of the first layer. Prior to MS measurement, the dried sample spot was washed twice. A drop of water (1 μL) was deposited over the sample spot and then blown off with a stream of air after a few seconds. This washing step removed salts present in the sample, minimizing salt-protein adduct formation and in turn improving the MS signal.

3.2.4 In-Capillary Digests Using a Nanoliter Chemistry Station. The nanoliter chemistry station[4-6] (nanochem station) is a custom-built sample-handling device

designed for nanoliter- and picoliter-volume samples. It was used to perform tryptic digestions on CE fractions and standard protein solutions, and then to deposit the digests for MALDI analysis. A schematic diagram of the nanochem station is shown in Figure 3.1. Briefly, the station consists of a 20 μm I.D. fused-silica capillary connected to a disposable syringe and mounted on a XYZ manipulator with video microscope observation. To minimize sample adsorption, the capillary is treated with a siliconizing agent (Glassclad-18, United Chemical Technologies, Bristol, PA) before use. MALDI targets or holders supporting disposable pipette tips containing samples and reagents are positioned on an XY stage close to the capillary tip.

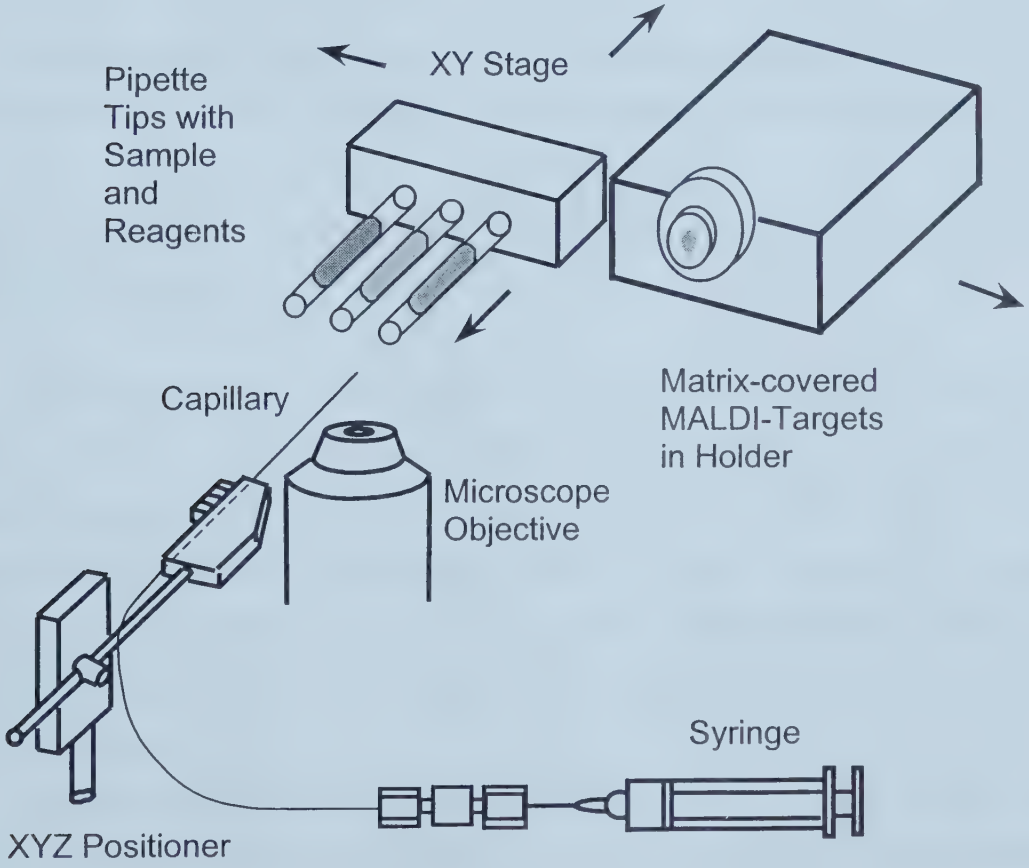

Figure 3.1 Schematic diagram of the nanoliter chemistry station. Figure is a modification of one taken from reference 6 courtesy of Bernd Keller.

The steps involved in performing digestions with the nanochem station are illustrated in Figure 3.2. First a 1-nL volume of sample was drawn into the capillary, and the solvent was evaporated with the aid of a nitrogen stream across the capillary tip. This step was repeated until 20 nL of sample was concentrated inside the capillary. The capillary was washed after every three drying steps to remove salts. This was done by drawing ~1 nL of Nanopure water in and out of the capillary several times. Buffered trypsin solution was then drawn into the capillary, the end was sealed against a Parafilm-covered holder, and the mixture was allowed to digest for 30 min.

For MALDI analysis, the target was prepared by depositing 1 μL of first-layer matrix solution (10 mg/mL HCCA in 20% methanol/acetone v/v) and allowing it to dry and form a thin layer of fine crystals. A 0.5-μL aliquot of second-layer matrix solution (33% methanol/water v/v saturated with HCCA) was then deposited on top of the first layer and allowed to dry. The matrix spot was washed twice with Nanopure water to remove any salts. Finally, 1 nL of second-layer matrix solution was drawn into the nanochem station capillary and deposited along with the digestion mixture onto the matrix-covered target.

3.2.5 On-Target Digests. Two types of MALDI targets were used for on-target digestion reactions: a plain polished stainless steel target and an AnchorChip™ target (Bruker Daltonics). The AnchorChip™ target is covered with a Teflon-like hydrophobic coating and an array of 400-μm hydrophilic spots or 'anchors'. When a μL-volume drop of sample or matrix solution is placed on one of the anchors, it shrinks to stay on the hydrophilic area as it dries. As a result, smaller sample spots are created than on a stainless steel target.

Prior to performing an on-target digestion, the MALDI target was thoroughly washed with acetone followed by methanol. The CE fraction or 3 μL of a standard protein solution was then deposited directly on the bare target 1 μL at a time and allowed to evaporate until almost completely dry. Next 1 μL of trypsin in 50 mM ammonium bicarbonate was added, the target covered to prevent evaporation and the mixture allowed to digest for 30 minutes. After allowing the drop to evaporate until almost completely dry once again, 0.5-1 μL of matrix solution was added and allowed to dry completely.

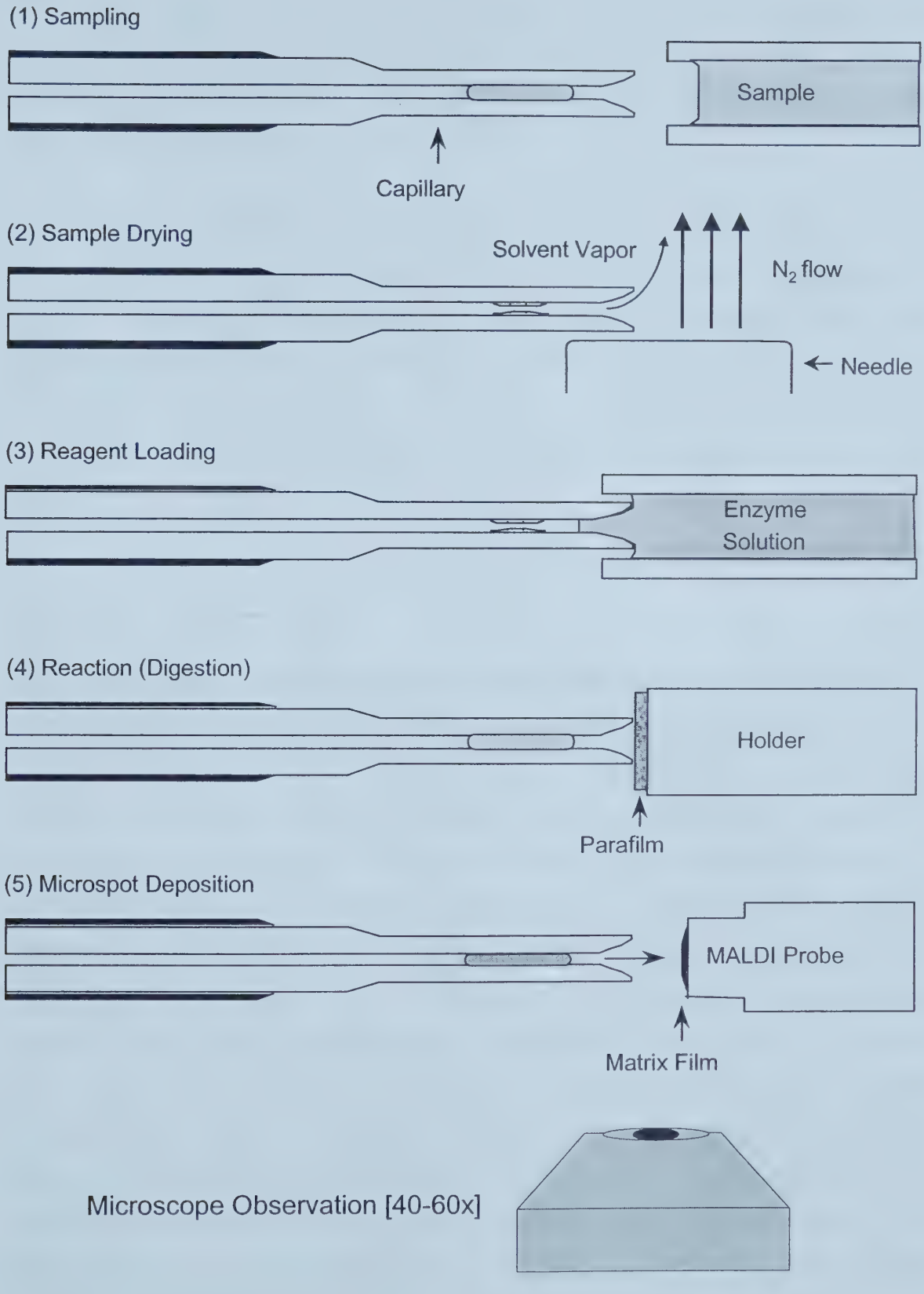

Figure 3.2 Sample preparation with the nanoliter chemistry station. Figure is a modification of one taken from reference 4 courtesy of Bernd Keller.

The matrix solution used for the stainless steel target was 33% MeOH/H2O (v/v) saturated with HCCA. The choice of matrix solution for the AnchorChip™ target is discussed in the Results section. Before MALDI analysis, the sample spot was washed twice with Nanopure water to remove salts.

3.2.6 MALDI MS Analysis. Mass spectral measurements of the in-capillary digests were conducted with a time-lag focusing MALDI-TOF mass spectrometer constructed in-house[7,8]. Mass spectral data acquisition and processing was performed with Hewlett Packard supporting software. The bulk and on-target digests were analysed using a Bruker Reflex III MALDI-TOF mass spectrometer (Bruker Daltonics) in reflectron mode. Data was reprocessed using IGOR Pro (Ver. 3.13, Wavemetrics Inc., Lake Oswego, OR) software.

3.3 Results and Discussion

3.3.1 Bulk Solution Digests. The first enzymatic digestion method considered for CE-collected proteins was bulk digestion, which is the simplest type to perform. The protein is simply mixed with a buffered enzyme solution in a vial and allowed to digest. This method works well with protein concentrations at or above the micromolar range, and is standard procedure in many labs[9]. To test if CE fractions could be digested in bulk, a 33 nM standard solution of cytochrome *c* (simulating a CE fraction with 100 fmol of cytochrome *c* in 3 μL of collection solution) was digested using a 5:1 enzyme/substrate molar ratio. Even though this ratio is considerably higher than the 1:10 ratio routinely used with higher protein concentrations, no peptide peaks were seen in the resulting digest mixture. This indicates that the trypsin concentration, which was only 75 nM after mixing with the sample and buffer solutions, was too low. The probability of trypsin coming into contact with cytochrome *c*, which was at an even lower concentration, would have been very small and therefore the rate of digestion would have been very low. In addition, some of the trypsin and cytochrome *c* would have adsorbed onto the vial wall, decreasing the chance of digestion even further.

The digestion was then repeated with a higher enzyme/substrate ratio of 12:1. Figure 3.3a shows the MALDI mass spectrum of the resulting digest. There are several trypsin autolysis peaks present in the spectrum, but no cytochrome *c* peptide peaks. The presence of autolysis peaks indicates that the trypsin was active, but the absence of cytochrome *c* peptide peaks is inconclusive. A further increase in the trypsin concentration might sufficiently raise the frequency of enzyme/substrate contact to generate enough cytochrome *c* peptides to be seen with MALDI MS. Accordingly, the enzyme/substrate ratio was increased to 24:1. This resulted in numerous trypsin autolysis peaks but only two cytochrome *c* peptide peaks (Figure 3.3b). When the trypsin concentration was doubled again, the autolysis peaks grew stronger but no additional cytochrome *c* peptide peaks were seen. Clearly, trypsin autolysis will always overpower cytochrome *c* digestion when so little substrate is present. A technique that includes a step to increase the substrate concentration is required for successful digestion of CE fractions.

3.3.2 In-Capillary Digests Using a Nanoliter Chemistry Station. The nanochem station has successfully been used to analyse sub-femtomole amounts of protein[5] and offers several advantages over other digestion techniques. For example, samples can easily be concentrated by performing multiple sample drying steps (steps (1) and (2) in Figure 3.2). In this work, a 20× concentration factor was achieved by drying 20 nL of sample in the capillary and then digesting it in 1 nL of trypsin. Furthermore, the multiple rinsing steps in between drying steps remove buffer salts and other contaminants. Since these contaminants are not concentrated along with the protein, there is less signal suppression during MALDI analysis. In addition, the use of only 20 nL of sample for each digestion means that several different enzymatic digestions, other chemical reactions, and direct MALDI-MS analysis can all be performed with a single CE fraction. The nanochem station also offers an advantage in MALDI sample deposition. Since the sample is deposited from the end of a small capillary instead of a micropipette tip, the resulting sample spot has a smaller area than samples deposited in the normal manner. The analyte is therefore less sparsely distributed and gives a stronger MALDI signal.

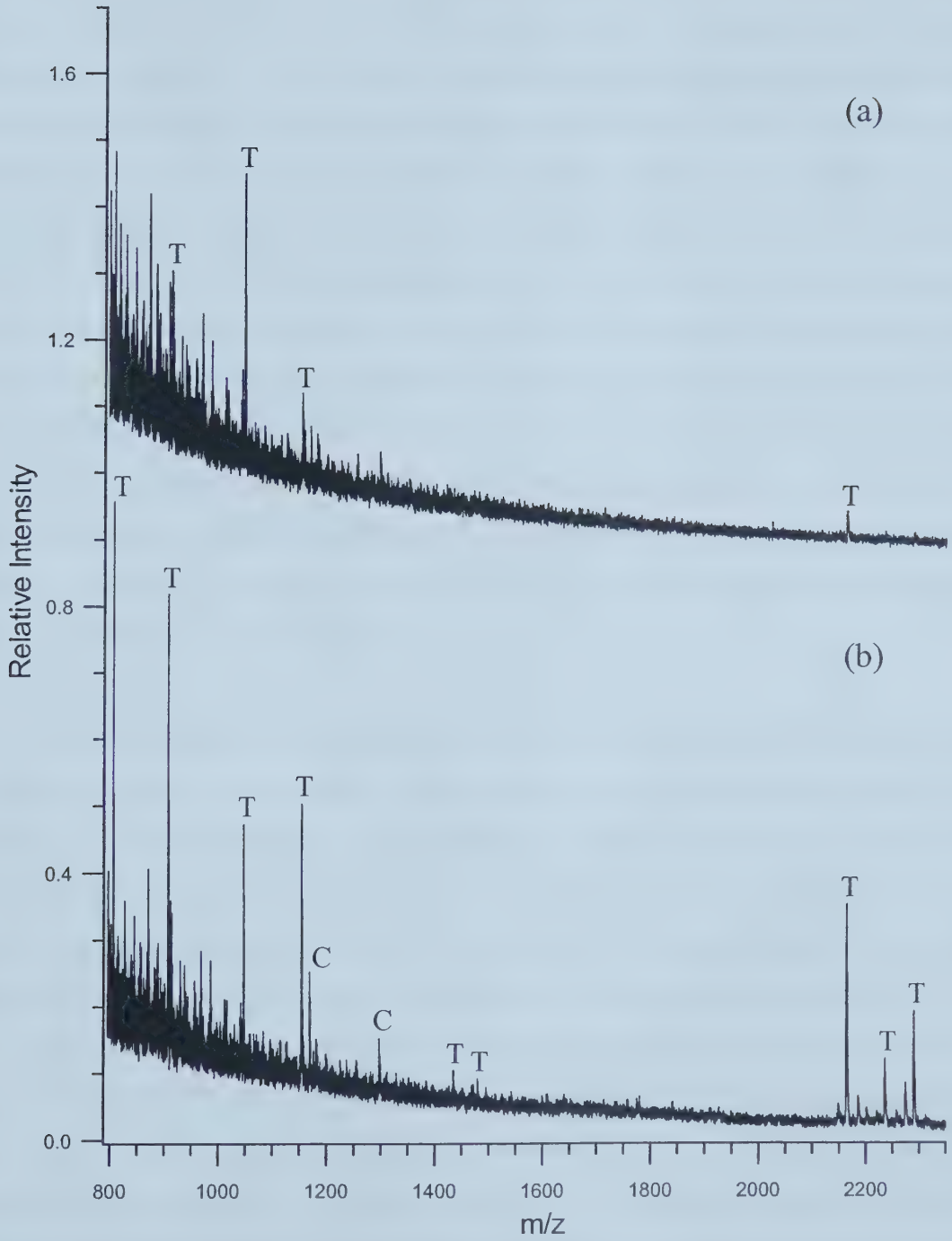

Figure 3.3 MALDI mass spectra of bulk digests of 33 nM cytochrome *c* standard solution using enzyme/substrate ratios of (a) 12:1 and (b) 24:1. Cytochrome *c* and trypsin autolysis peaks are labelled C and T respectively.

The many benefits of the nanochem station make it a promising candidate for digesting CE-collected proteins, and initial in-capillary digestions were successful. For example, Figure 3.4a shows the MALDI mass spectrum of the digest of a CE fraction containing 200 nM BSA. Fourteen BSA peptide peaks can be identified that cover 35% of the amino acid sequence. However, at protein concentrations closer to the CE/MALDI MS whole-protein detection limits reported in Chapter 2 (23 fmol for cytochrome *c*), in-capillary digestions of both standards and CE fractions were unsuccessful. For example, Figure 3.4b shows the MALDI mass spectrum of the digest of a 33 nM cytochrome *c* solution. Several trypsin autolysis peaks and many unidentified low-mass peaks are visible, but only a few very weak cytochrome *c* peptide peaks are present. Digests of CE fractions containing the same amount of cytochrome *c* gave similar results. Even with the multiple drying steps prior to digestion, the substrate concentration is too low to allow sufficient digestion for protein identification. The goal of this work is to be able to digest proteins at even lower concentrations than 33 nM, so another digestion method with a higher concentration factor is required.

3.3.3 On-Target Digests. As discussed in Section 3.2.5, AnchorChip™ MALDI targets are designed with a hydrophobic coating and 400-μm hydrophilic 'anchors' to concentrate spotted samples into a smaller area than conventional uncoated targets. The increased spatial density of the analyte allows more analyte to be desorbed with each laser shot, thereby increasing the MALDI sensitivity. On-target digestions were first attempted on AnchorChip™ targets to take advantage of this increased sensitivity.

Unfortunately, the AnchorChip™ targets did not perform as well as expected. In general, AnchorChip™ digests gave peptide signals that were weak and originated mainly from trypsin rather than the protein of interest. The results grew worse with lowered protein concentration. The main reason for the unsatisfactory MALDI results was poor matrix crystallization. Previous experiments in our research group have shown that the best MALDI results are obtained with a dense layer of small, uniform matrix crystals.[10] All of the samples on the AnchorChips™ formed large isolated crystals, sometimes even outside of the anchor spots on the hydrophobic coating. In an attempt to

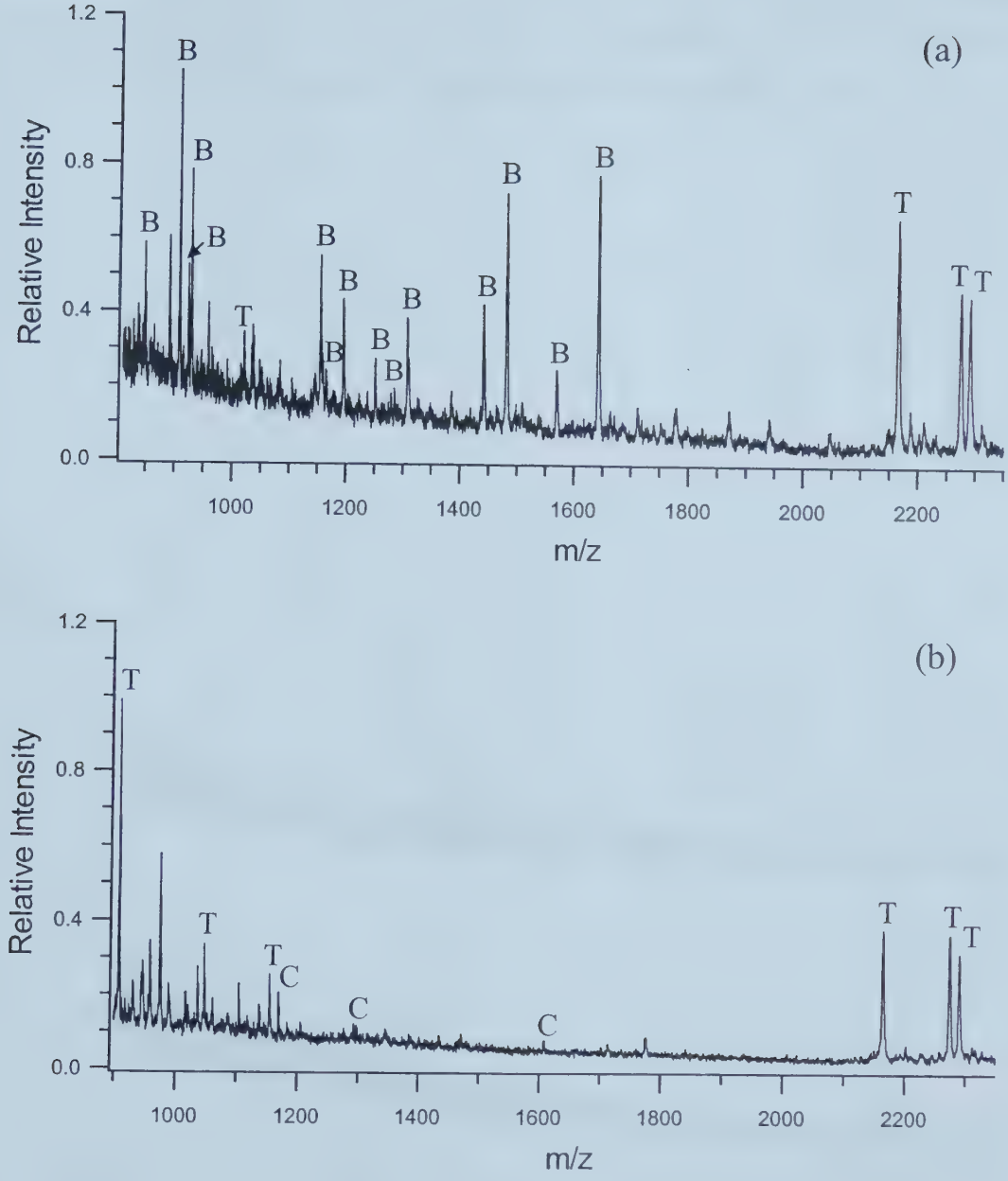

Figure 3.4 MALDI mass spectra of in-capillary digests of (a) a CE fraction containing 200 nM BSA and (b) a standard solution containing 33 nM cytochrome *c*. BSA, cytochrome *c* and trypsin autolysis peaks are labelled B, C and T respectively.

improve the crystallization, the dried sample spots were washed with 1 μL of 10% formic acid in water and recrystallized with 0.5 μL of 6:3:1 ethanol/acetone/10% formic acid (v/v/v). In some cases, washing and recrystallization was helpful. For example, Figure 3.5a shows the MALDI mass spectrum of an on-target digest of 100 fmol cytochrome *c*. The signal intensity is strong and eleven cytochrome *c* peptide peaks can be identified. In other cases, however, washing and recrystallization had no beneficial effects.

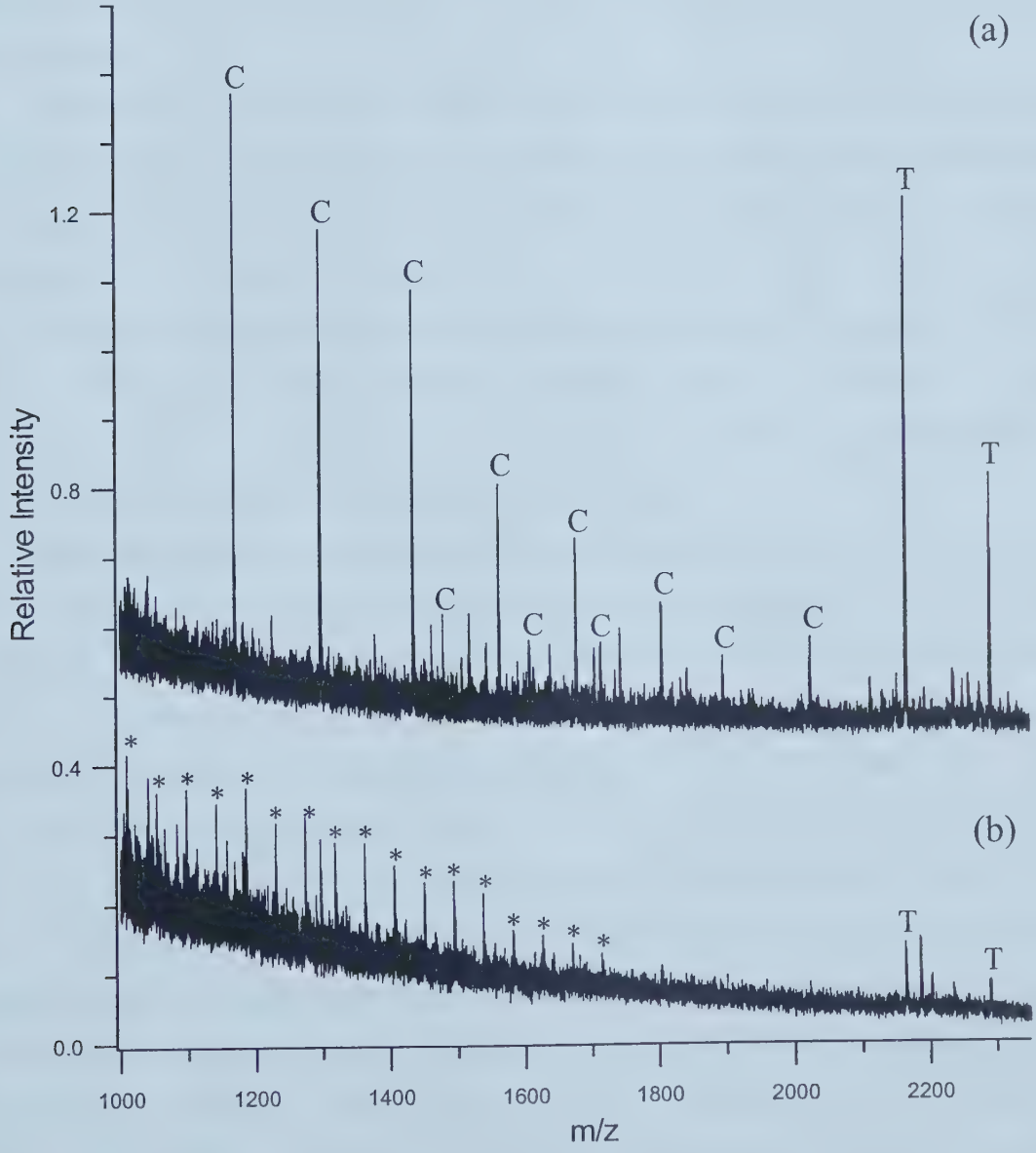

Figure 3.5 MALDI mass spectra of on-target AnchorChip™ digests of standard solutions containing (a) 100 fmol and (b) 20 fmol cytochrome *c*. Cytochrome *c*, PEG and trypsin autolysis peaks are labelled C, * and T respectively.

A number of other variations in sample preparation were also tried in an effort to improve crystallization. Different HCCA concentrations ranging from 0.2 g/L to saturation and solvent mixtures containing 20-100% methanol or acetonitrile in 0.1% TFA were evaluated for the matrix solution. A solution of 0.3 g/L HCCA in 100% methanol was found to be the best of the variations tried, but it still produced large crystals. Drying the samples faster by applying a stream of hot air or by heating the target was also tried in an attempt to prevent the crystals from becoming too large, but with no success.

Another major problem that appeared in most of the AnchorChip™ digests was poly(ethylene glycol) (PEG) contamination, especially at lower protein concentrations. For example, Figure 3.5b shows the mass spectrum of an on-target digest of 20 fmol cytochrome *c*. Some trypsin autolysis peaks are seen, but the 1000-2000 Da region where cytochrome *c* peptide peaks are normally seen shows only PEG peaks, identified by the 44 Da mass difference between successive peaks. The issue of PEG contamination will be discussed later in more detail. Overall, the performance of AnchorChips™ for on-target digestions was disappointing.

Stainless steel targets provided much better digestion results. Figure 3.6a shows the MALDI mass spectrum of a standard solution digest containing 100 fmol cytochrome *c*. Nineteen cytochrome *c* peptide peaks are identified that cover 79% of the amino acid sequence. Even when the amount of protein is reduced to 20 fmol, fifteen peaks can still be identified, giving 60% sequence coverage (Figure 3.6b). CE fraction digests were comparable to the standard protein digests at the 100- and 50-fmol levels. As shown in Figure 3.7a, the digest of a CE fraction containing 92 fmol cytochrome *c* exhibits the same peptides as the 100 fmol standard shown in Figure 3.6a. As with the AnchorChip™ digests, however, PEG contamination became a problem at lower concentrations. Figure 3.7b shows the mass spectrum of a 23 fmol cytochrome *c* CE fraction digest. A few trypsin autolysis peaks are visible, but many of the cytochrome *c* peptide peaks are overshadowed by PEG. While some cytochrome *c* peaks can be manually detected amidst the PEG peaks by comparing the peak masses to a theoretical digest list, it would be difficult for computer software to isolate them for MS/MS analysis. For example, Figure 3.8 shows an expanded view of two cytochrome *c* peaks from Figure 3.7b. The

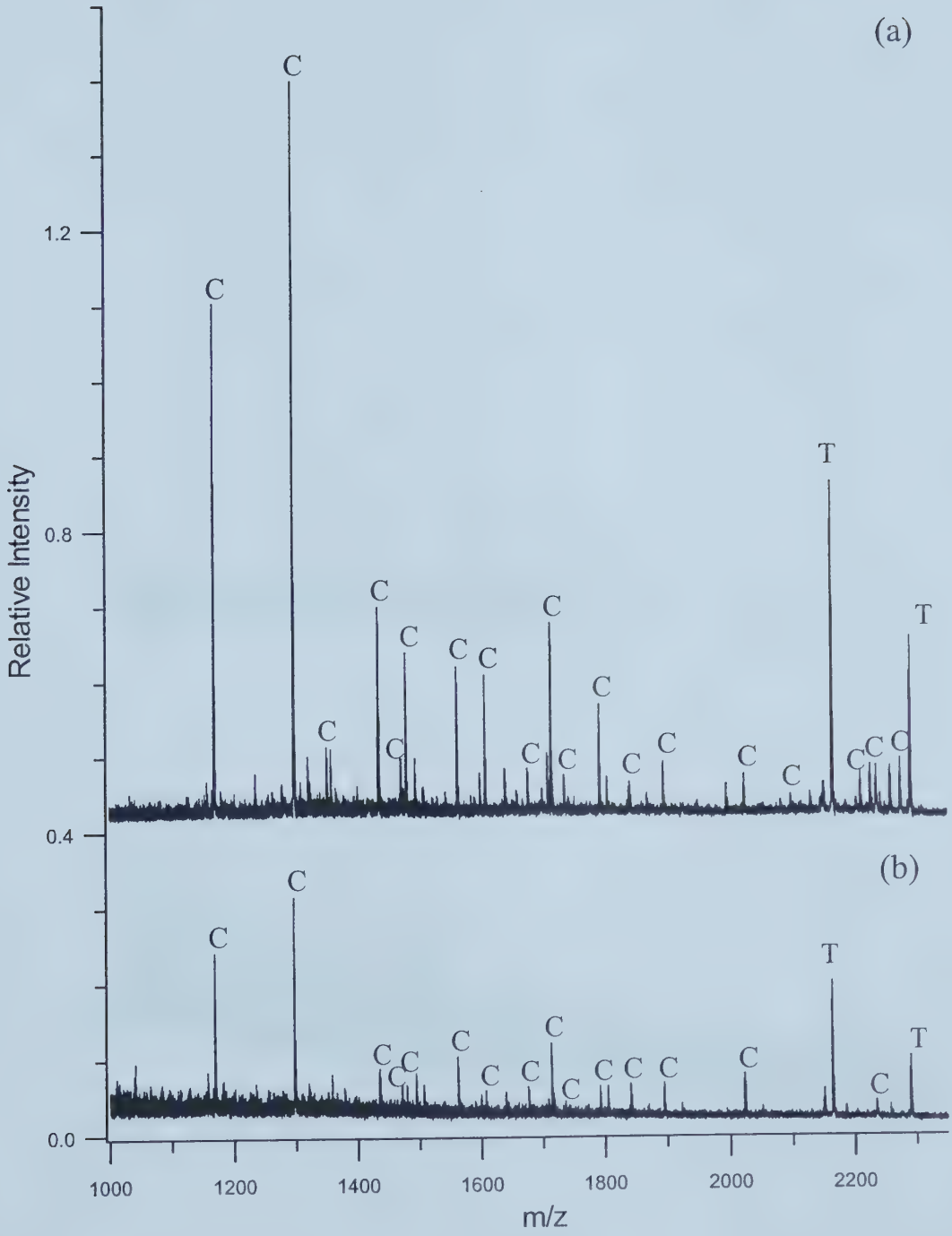

Figure 3.6 MALDI mass spectra of stainless steel on-target digests of standard solutions containing (a) 100 fmol and (b) 20 fmol cytochrome *c*. Cytochrome *c* and trypsin autolysis peaks are labelled C and T respectively.

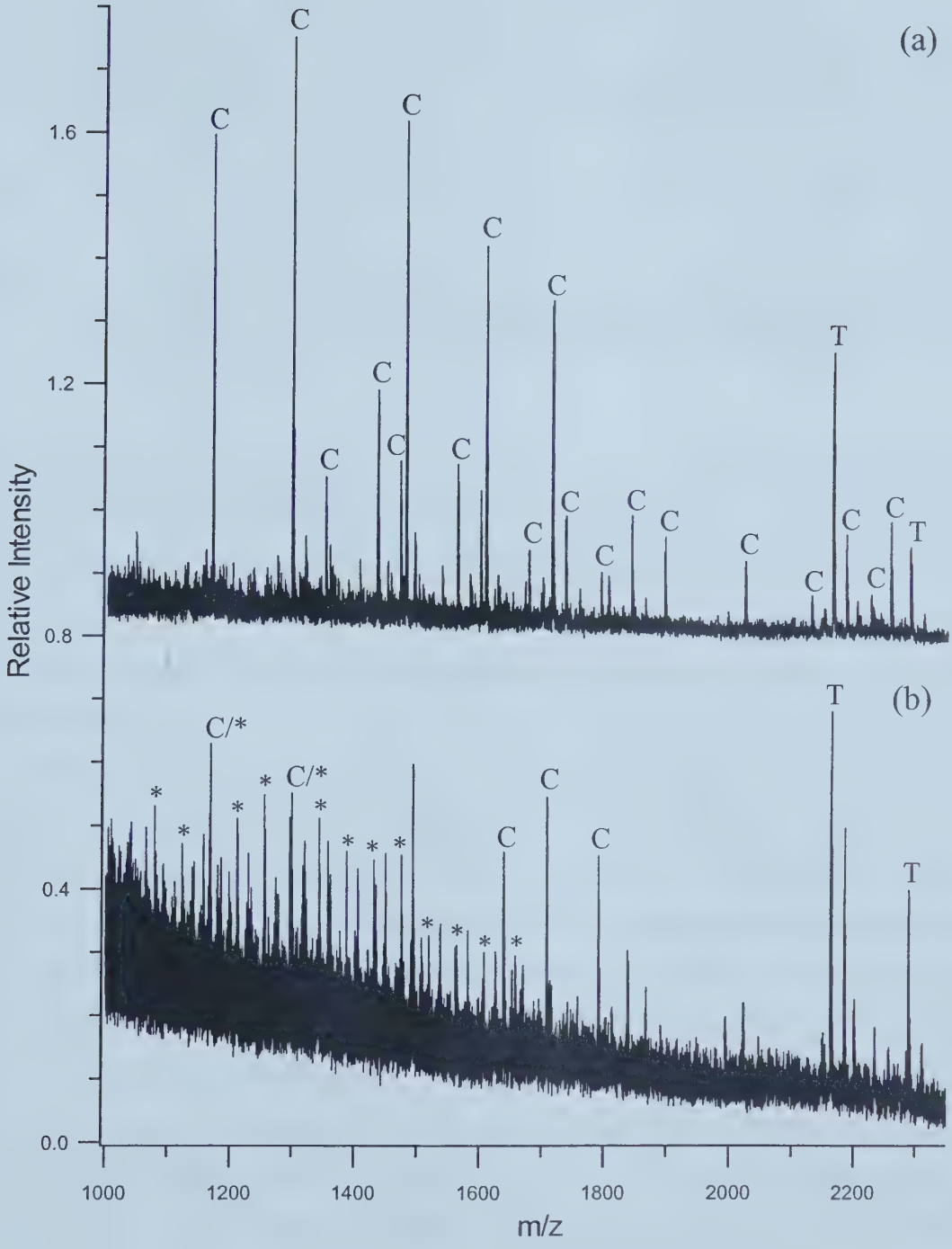

Figure 3.7 MALDI mass spectra of stainless steel on-target digests of CE fractions containing (a) 92 fmol and (b) 23 fmol cytochrome *c*. Cytochrome *c*, PEG and trypsin autolysis peaks are labelled C, * and T respectively.

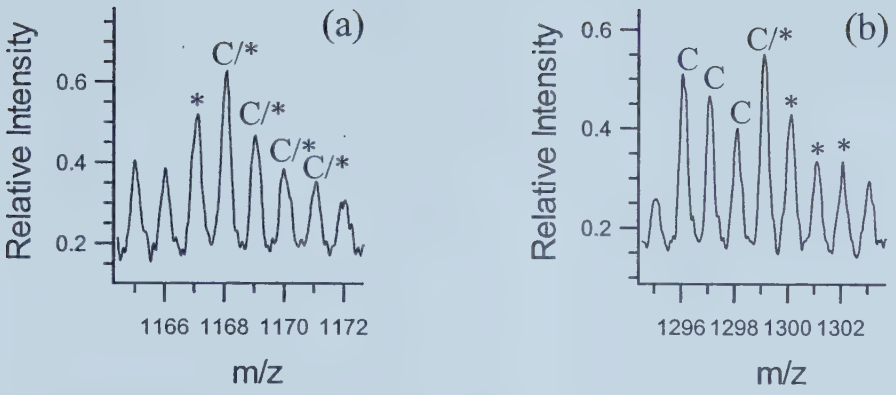

Figure 3.8 Expanded view of two cytochrome *c* peptide peaks and overlapping PEG contamination from Figure 3.7b. Cytochrome *c* and PEG peaks are labelled C and * respectively.

1168 Da cytochrome *c* peak (Figure 3.8a) can only be distinguished from the 1167 Da PEG peak by the shift in the isotopic distribution, which is higher at 1168 than 1167 Da. The 1296 Da cytochrome *c* peak (Figure 3.8b) is separated a little more from the 1299 Da PEG peak, but it is still too close for automated detection. With this degree of PEG interference, it would be impossible to perform *de novo* sequencing in order to identify an unknown protein.

When investigated more closely, two recurring PEG distributions were identified in the MALDI mass spectra of the on-target digests. The PEG distribution that gave the stronger MALDI signal is shown in Figure 3.9a. It was seen in CE fraction digests and in the DDAB solution used to coat the capillary, but not in standard digests, in undigested CE fractions or in matrix or water blanks. The fact that it was seen in CE fraction digests but not in standard digests indicates that it does not come from the buffered trypsin solution (trypsin in 50 mM NH_4HCO_3) used for digestion. This raises the question of how it could appear in digested CE fractions but not undigested fractions. If it is not from the trypsin solution itself, its appearance could be induced by the increased pH during digestion. This theory was tested by 'digesting' a CE blank (a CE fraction containing no protein) with 50 mM NH_4HCO_3 but no trypsin. When analysed by MALDI, PEG did in fact appear. All of these indications suggest that this main type of PEG comes from somewhere in the CE process, possibly from the DDAB capillary coating, and is brought out by the increased pH during digestion.

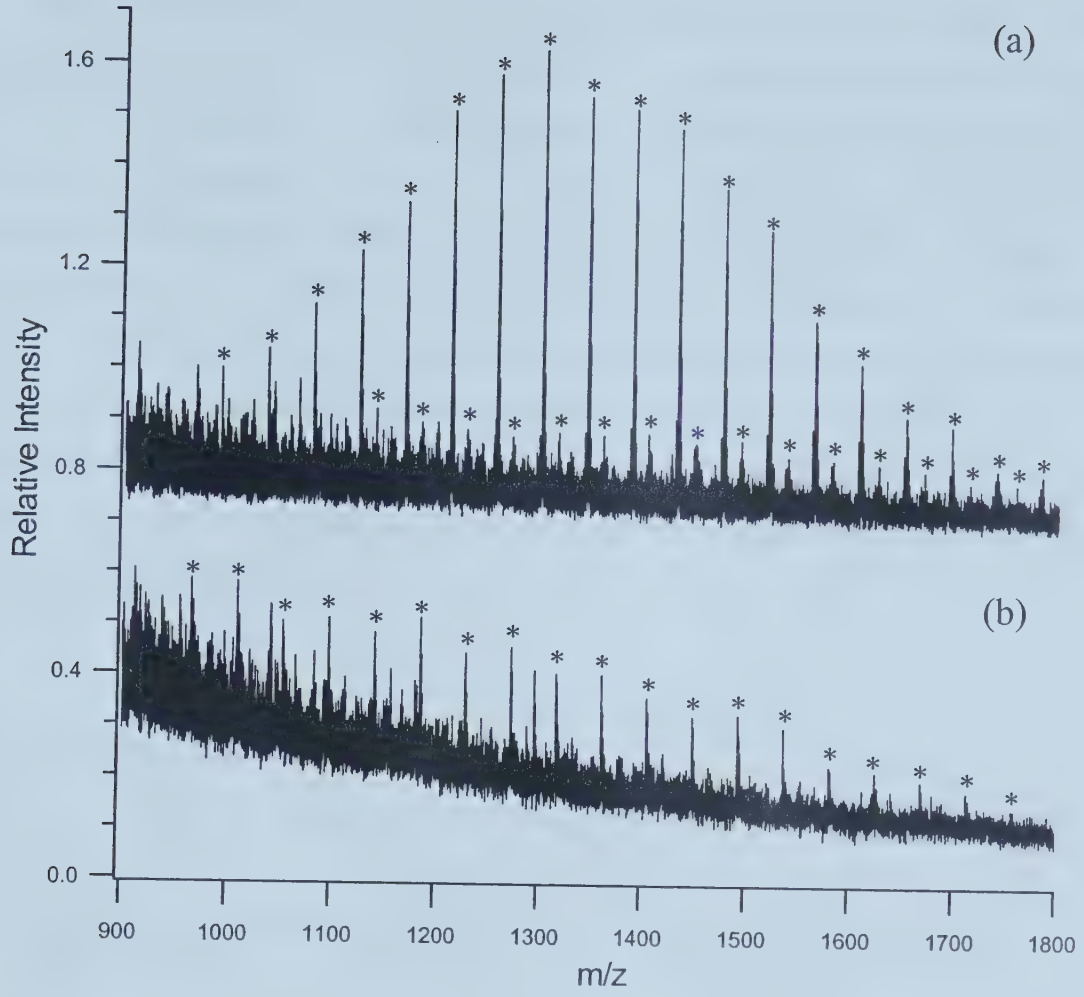

Figure 3.9 MALDI mass spectra of PEG contamination in (a) a digested CE fraction and (b) a digested standard solution. PEG peaks are labelled * .

The second recurring PEG distribution, which gave a much weaker MALDI signal, is shown in Figure 3.9b. It was seen in all of the digested samples, both standards and CE fractions, though not in undigested CE fractions or in matrix or water blanks. This PEG is most likely a general 'background' contaminant that is brought out by the higher pH during digestion like the main type of PEG. It could come from any number of possible sources such as the MALDI matrix, the trypsin and standard proteins used, solvents, plastic bottles and vials, and innumerable other laboratory accessories. PEG contamination has, in fact, appeared sporadically during MALDI analysis by various members of our research group, and its source has never been isolated.

The main conclusion derived from these on-target digestion experiments is that peptide mapping cannot be successfully performed on low-concentration CE fractions until the PEG contamination is eliminated. In fact, any digestion method that involves a non-specific concentration step will be unsuccessful because the PEG will be concentrated along with the proteins. Unfortunately, PEG is a difficult contaminant to eliminate because it is such a common polymer. In this experiment, for example, attempts at PEG reduction such as using glass vials instead of plastic, using fresh reagent solutions each day, and using freshly deionized water and newly-opened bottles of solvents were all ineffective. If the PEG contamination cannot be eliminated, a sample preparation method that concentrates proteins but not contaminants like PEG will have to be used prior to digestion.

3.4 Literature Cited

(1) James, P., Ed. *Proteome Research: Mass Spectrometry*; Springer-Verlag: Berlin, 2001.

(2) Palmer, T. *Understanding Enzymes*; 4th ed.; Prentice Hall/Ellis Horwood: London, 1995.

(3) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999**, *71*, 1087-1091.

(4) Whittal, R. M.; Keller, B. O.; Li, L. *Anal. Chem.* **1998**, *70*, 5344-5347.

(5) Keller, B. O. *Ph.D. thesis*; The University of Alberta: Edmonton, 2001.

(6) Keller, B. O.; Li, L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1055-1063.

(7) Whittal, R. M.; Li, L. *Anal. Chem.* **1995**, *67*, 1950-1954.

(8) Whittal, R. M.; Russon, L. M.; Weinberger, S. R.; Li, L. *Anal. Chem.* **1997**, *69*, 2147-2153.

(9) Coligan, J. E.; Dunn, B. M.; Ploegh, H. L.; Speicher, D. W.; Wingfield, P. T., Eds. *Current Protocols in Protein Science*; Supplement 20 ed.; John Wiley & Sons, Inc.: New York, 2000; Vol. 2.

(10) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996**, *68*, 2494-2500.

Chapter 4

Capillary Zone Electrophoresis Combined with Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for the Analysis of Complex Protein Mixtures from Cell Lysates of *Escherichia coli* [a]

4.1 Introduction

At present, proteome display is most commonly performed with polyacrylamide gel electrophoresis (PAGE). While gel electrophoresis is very powerful, it is a slow and laborious technique. It can easily take two days to cast a gel, run it, stain it, excise the sample spots, destain the gel pieces, perform in-gel digestion if desired, and finally extract the proteins or peptides for MS analysis. Furthermore, low-concentration proteins are difficult to recover from gels and often cannot be efficiently digested in-gel. Research is therefore driven towards the development of alternative techniques with high sample throughput and MS compatibility.

A number of other separation techniques coupled with MS have been reported for the analysis of bacterial proteomes. High-performance liquid chromatography (HPLC) has been combined with both electrospray ionization (ESI) and MALDI MS[1-3] with favourable results. However, HPLC separations still tend to be time-consuming and often require high sample loading, which may not be feasible when dealing with limited amounts of biological sample. Capillary electrophoresis (CE) techniques, on the other hand, offer multiplex capability, have fast analysis times, and can handle trace amounts of samples with high separation efficiency. Capillary isoelectric focusing (CIEF) has been coupled on-line with various MS methods[4-7] for bacterial analysis, but the required on-line interfaces are temperamental and can restrict separation conditions. Our method combining CZE off-line with MALDI MS is simple, sensitive, and fast. It also allows the options of subjecting the proteins collected in the CE fractions to multiple chemical or

[a] A version of this chapter has been submitted for publication: Ken K.-C. Yeung, Andrea G. Kiceniuk and Liang Li "Capillary Zone Electrophoresis Combined with Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for the Analysis of Complex Protein Mixtures from Cell Lysates of *Escherichia coli*"

enzymatic reactions as discussed in Chapter 3, and storing any unused portions for future analysis.

The two preceding chapters have dealt with the CZE analysis of standard proteins such as lysozyme and cytochrome *c*. The analysis of real world samples such as cell lysates is a much more challenging problem. Even though they are relatively simple organisms, *E. coli* bacteria are known to express hundreds of proteins. Not all of these proteins are recovered during extraction, but *E. coli* cell lysates still contain a complex mixture of proteins that exhibit a wide range of p*I*s, molecular masses, and hydrophobicities. In addition, there are other background substances present including salts, buffers, and other non-protein molecules. This chapter discusses the application and refinement of our CZE/MALDI-MS method for the analysis of these complex samples.

4.2 Experimental

4.2.1 Materials. All solutions were prepared in Nanopure 18MΩ water (Barnstead). Buffers were prepared from reagent grade orthophosphoric acid (Mallinckrodt, Paris, KY), and the pH was adjusted with reagent grade sodium hydroxide (Fisher Scientific, Fair Lawn, NJ). Hydrochloric acid used in fraction collection was purchased from Anachemia (Montreal, PQ). Didodecyldimethylammonium bromide (DDAB), α-cyano-4-hydroxycinnamic acid (HCCA), spectrophotometric grade trifluoroacetic acid (TFA) and lyophilized *Escherichia coli* cells (ATCC 9637, strain W) were purchased from Sigma-Aldrich Canada (Oakville, ON). HPLC grade acetonitrile was purchased from Fisher. All reagents were used as received except for the HCCA, which was purified by recrystallization from ethanol prior to use.

4.2.2 *Escherichia coli* Lysate Preparation. Extraction of proteins from the *E. coli* cells was performed by solvent suspension in 0.1 % v/v aqueous TFA. Such acid extractions are selective toward low molecular mass proteins (<15 000 Da)[8]. When the lysate was to be analysed directly by CE, 125 mg of lyophilized cells were suspended in 7 mL of 0.1 % TFA, placed in an ultrasonic bath (Cole-Parmer, model 8851) for 5 min and then vortexed

for 5-10 min. The supernatant was isolated by centrifugation (12 000 g), concentrated by evaporation (Speed-Vac, Eppendorf Vacufuge™) to *ca.* 700 μL (10-fold), then further concentrated with a 3000 Da molecular mass cutoff filter (Microcon YM-3, Millipore Corp., Bedford, MA) to 125 μL prior to CZE injection. When the lysate was to be pre-separated by liquid chromatography prior to CE, 100 mg of *E. coli* was extracted for each LC separation. The supernatant was concentrated by evaporation to *ca.* 500 μL and then loaded directly onto the column.

4.2.3 Protein Separation by CZE. All CZE separations and fraction collections were performed with an Agilent ^{3D}CE instrument (Palo Alto, CA) equipped with an ultraviolet-visible diode array detector. Fused silica capillaries (Polymicro Technologies, Phoenix, AZ) with an inner diameter of 50 μm, an outer diameter of 365 μm, and a total length of 30.5 cm (22.3 cm to the detector) were used. The applied voltage used during electrophoresis was –6 kV. The capillary was thermostated at 25°C in all experiments. Data acquisition and instrument control was performed using ChemStation software (Rev. A.06.03, Agilent) on a Pentium based microcomputer. Direct UV absorbance detection was performed at 192 nm.

New capillaries were used for each protein sample to avoid possible cross contamination. Prior to separation, each capillary was preconditioned with a high-pressure (100 kPa) flush of NaOH for 5 min and distilled water for 1 min. Initial treatment of the capillary surface with DDAB was performed by a 5-min rinse (100 kPa) with 0.1 mM DDAB in water, followed by a 1-min rinse (100 kPa) with the electrophoretic buffer[9]. A pH 4.0 sodium phosphate buffer (25 mM) was used in all experiments. Between subsequent separations, the capillary was rinsed (100 kPa) with 0.1 mM DDAB for 2 min and electrophoretic buffer for 1 min. Sample injection was performed hydrodynamically at 5 kPa for 1.0 to 10.0 seconds, depending on the sample concentration.

4.2.4 CZE Fraction Collection. Collection of an analyte band at the outlet of the capillary was performed electrokinetically into polypropylene micro inserts (Agilent, Part no. 5182-0549). As described in Chapter 2, the fraction collection timing was calculated

based on the analyte's migration time and the ratio between the total and effective lengths of the capillary (Equation 2.1). A dilute (10 ppm) solution of HCl was placed in the micro insert to maintain electrical contact between the capillary and electrode. The use of HCl solution instead of electrophoretic buffer minimized the amount of sodium ions in the fraction and improved the MALDI signal (see Chapter 2)[10,11].

To reduce dilution during fraction collection, the volume of dilute HCl used was minimized. The smallest amount required was 2 μL, which resulted in a solution depth of ~0.75 mm at the bottom of the micro-insert. However, the 2-μL volume did not result in successful collection in every case, so a 4-μL volume was used in most cases to ensure electrical contact between the electrode and the capillary outlet. The top 5 mm of the micro-insert was trimmed off to raise the position of the insert and thus bring the electrode and capillary outlet closer to the bottom of the micro-insert.

4.2.5 Pre-Separation by LC. A preparative-scale crude separation was performed by liquid chromatography in some cases prior to CZE. A laboratory-packed column was prepared using reverse-phase C8 particles with 20-30 μm diameters (208TP2030, Vydac). Approximately 0.5 g of C8 particles were slurry-packed into a Pasteur pipette with a small piece of glass wool used as an outlet frit. This yielded a chromatographic bed of *ca.* 0.5 cm diameter by 4.5 cm long. The concentrated *E. coli* lysate solution was loaded directly onto the column and washed with 1.0 mL of 0.05% v/v aqueous TFA. Then the sample was eluted with consecutive 0.5 mL portions of mobile phase with increasing acetonitrile content (v/v): 5 portions of 5%, 5 portions of 10%, 3 portions of 20%, 2 portions of 30%, 2 portions of 40%, 1 portion of 50% and 1 portion of 60% in 0.05% (v/v) TFA. Each fraction was collected in a separate vial for a total of nineteen fractions in all.

4.2.6 MALDI Sample Preparation and MS Analysis. A two-layer sample deposition procedure was used for proteins extracted from *E. coli*[12]. The MALDI target was prepared by depositing 1 μL of the first-layer matrix solution (10 mg/mL HCCA in 20% methanol/acetone v/v) and allowing it to dry and form a thin layer of fine crystals. The sample was then mixed 1:1 with FPW matrix solution (1:2:3 formic acid/isopropanol/

water saturated with HCCA), and 0.25 μL of the mixture was deposited on top of the first layer. Experience in our research group has shown that FPW matrix increases the MS sensitivity of proteins extracted from bacterial cells. The use of high concentration acids has also been reported by others to improve MS signals in the presence of salts[13].

Prior to MS measurement, the dried sample spot was washed twice. A drop of water (1 μL) was deposited over the sample spot and then removed with a Kimwipe after 3 seconds. This washing step removed salts present in the sample, minimizing salt-protein adduct formation and in turn improving the MS signal.

Mass spectral measurements were conducted with a time-lag focusing MALDI-TOF mass spectrometer constructed in-house[14,15]. Mass spectral data acquisition and processing was performed with Hewlett Packard supporting software. Data was reprocessed using IGOR Pro (Ver. 3.13, Wavemetrics Inc., Lake Oswego, OR) software.

4.3 Results and Discussion

4.3.1 CZE and MALDI MS. Figure 4.1A shows the electropherogram resulting from a 1-second injection of raw (without LC pre-separation) *E. coli* cell extract. Partially resolved peaks and shoulders are observed from 2 to 8 minutes. The acid (TFA) extraction used in this experiment is selective towards proteins with neutral or alkaline isoelectric points. These basic proteins are cationically charged under the CZE conditions (pH 4.0), hence their electrophoretic mobility is cathodic. Capillaries treated with DDAB have cationic inner walls, so a fast anodic electroosmotic flow (EOF) is generated. As a result, the cationic proteins migrate opposite to the EOF. The magnitude of the reversed EOF is large (6.3×10^{-4} cm^2/Vs), and it is sufficient to carry the counter-migrating proteins toward the anode. Hence reversed polarity is used and the migration order is: anions, neutral molecules, and finally cations.

The first peak observed at 2.5 minutes in Figure 4.1A is trifluoroacetate. Trifluoroacetate is a fast moving anion, thus it arrives at the detector first under reversed polarity. The observed triangular peak shape, due to peak fronting, is characteristic of electrodispersion. Electrodispersion results from a difference in mobility between the analyte (trifluoroacetate) and the background co-ion (phosphate)[16]. Aside from the TFA

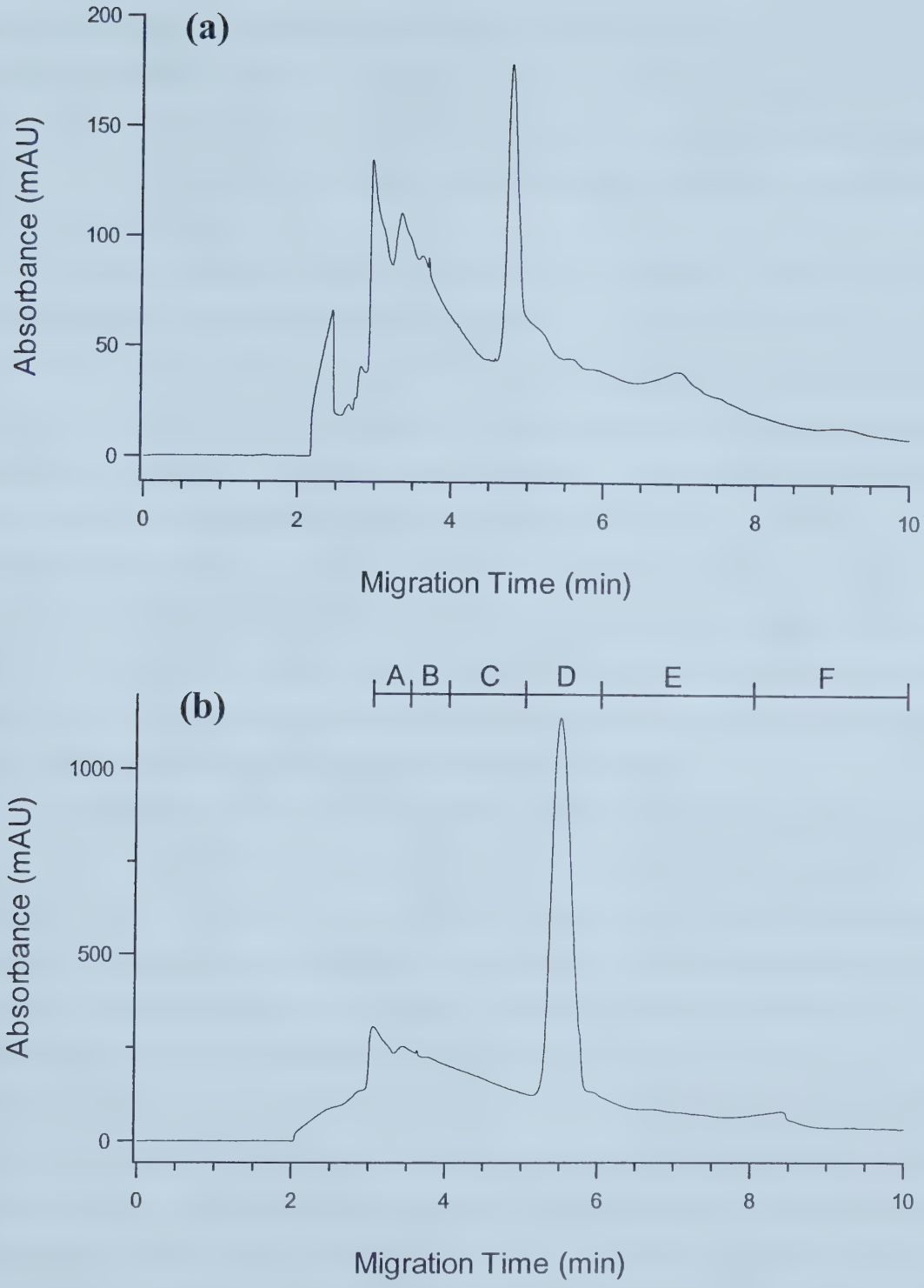

Figure 4.1 Electropherograms of (a) 1-second and (b) 5-second injections of raw *E. coli* extract. Separation is performed at -6 kV in a 50 μm I.D., 22.3 cm (effective)/30.5 cm (total) DDAB-coated capillary with 25 mM sodium phosphate buffer, pH 4.0. Fraction collection intervals for the 5-second injection are indicated by the lettered scale above the electropherogram.

peak, strong signals are mainly detected between 3 and 5 minutes. Injection of a neutral marker (acetonitrile) resulted in a migration time of 3 minutes. This result indicates that the signals observed around this region are from neutral or slightly charged molecules. The signals beyond 3 minutes are from cationic molecules, with faster cations detected at later migration times.

The peak efficiency observed in Figure 4.1 is rather poor. Previously reported peak efficiencies from DDAB-treated capillaries were greater than 500,000 plates/m[9]. The reported values, however, were obtained from the analysis of standard proteins such as α-chymotrypsinogen A, ribonuclease A, cytochrome *c* and lysozyme using optimal separation conditions for minimized band broadening. These results could be readily reproduced with our instrument using the same conditions. However, in this work the proteins are from cell lysates of *E. coli*. Proteins of various molecular masses, p*I*s and concentrations are extracted in TFA. Hundreds of protein species together with salts and TFA are expected to be present in the sample. Thus band broadening cannot be easily controlled. The combination of significant band broadening and incomplete resolution result in the broad, overlapping peaks observed in Figure 4.1A.

To investigate the degree of separation achieved under the conditions used in Figure 4.1A, MALDI-MS measurements were performed on fractions collected at the capillary outlet. Unfortunately, a useful MS signal was not detected from any of the fractions. With standard proteins such as cytochrome *c*, our CZE fractionation procedure allows the MALDI detection of less than 25 fmol of protein (see Chapter 2)[17]. It is therefore likely that the individual *E. coli* fractions contain very little protein. In an attempt to increase the MALDI signals, the volume of HCl solution used for fraction collection was reduced from 4 μL to 2 μL in order to decrease analyte dilution. With this change, however, MALDI protein signals were still not detected. Since the cell extract sample could not be further concentrated without precipitation, another option was to increase the analyte loading to CZE by increasing the injection volume. In this case, the injection volume was increased from 2.8 nL (1.0-second injection) to 14.1 nL (5.0-second injection). The resulting electropherogram is shown in Figure 4.1B.

Figure 4.1B shows that there is an overall increase in the UV intensity with the larger sample loading. However, the efficiency and resolution has decreased as a result

of the longer injection time. A 5.0-second injection time corresponds to an injection plug length of 7.2 mm. To avoid injection band broadening, the injection plug length should be limited to 1 mm for the size and length of the capillary used in this case[18]. However, the increased injection was necessary to deliver a sufficient amount of protein for subsequent MS measurement. Six fractions were collected as indicated in Figure 4.1B and their corresponding MALDI mass spectra are shown in Figures 4.2A-F. The mass spectrum of the original sample before CE separation is shown in Figure 4.2O.

As Figure 4.1B shows, fractions A to C are from the migration region of neutral or slightly charged proteins (3 to 5 minutes), whereas fractions D to F are from the region of the cationic proteins (5 to 10 minutes). One major protein (2579 Da) is observed from fractions A to C, but the majority of proteins are found in fractions D to F. There is considerable overlap in the proteins found in these later fractions. However, there are also marked differences in the relative intensities of the common peaks in each fraction. For example, the 7746 Da protein peak is much stronger in fraction E than in fractions D or F. In addition, there are some proteins that appear in only one fraction, such as the proteins at 10534 Da in fraction D and 8225 Da in fraction F. These differences show that there are variations in the protein compositions of the different fractions, indicating that CE achieves partial separation of the proteins in the cationic region (5 to 10 minutes).

Eight *E. coli* proteins are identified from Figures 4.2A-F, as listed in the first column of Table 4.1. The proteins were identified by mass comparison (0.05% mass accuracy) with the SWISS-PROT and TrEMBL *E. coli* proteome databases searched using the Sequence Retrieval System (http://ca.expasy.org/srs5) and TagIdent (http://ca.expasy.org/tools/tagident.html). The number of proteins identified from the CZE fractions is the same as those identified from the MALDI spectra of the original *E. coli* extract (Figure 4.2O). In other words, the CZE separation does not provide any new information. CZE alone is unable to isolate the less abundant proteins and reduce the signal suppression during MALDI-MS measurement so that more proteins can be identified. The protein mixture in the raw *E. coli* lysate is too complex for the 10-minute CZE separation. In order to increase the number of proteins identified, the use of liquid chromatography was investigated as a sample cleanup or pre-separation method for the raw extract prior to CZE.

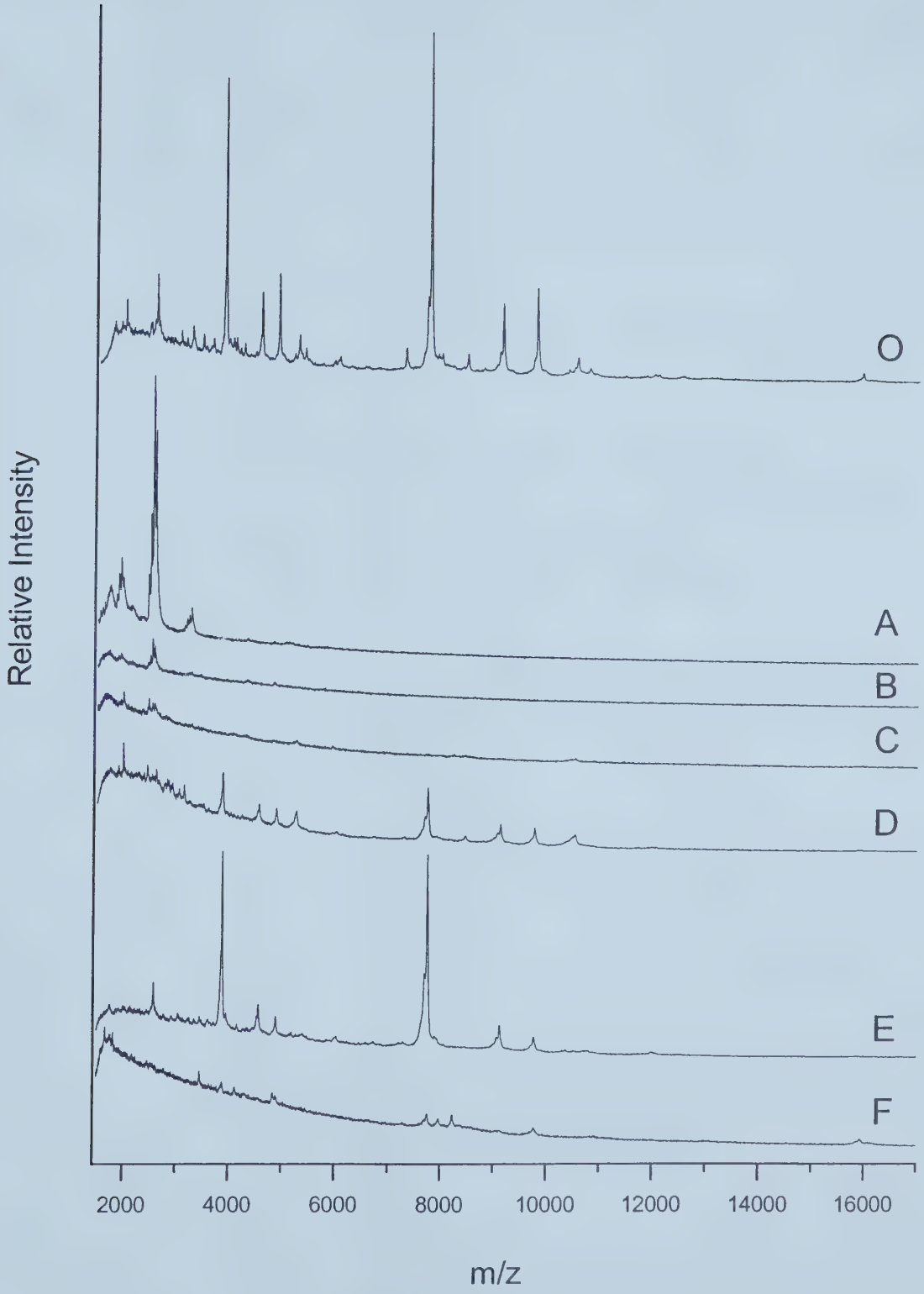

Figure 4.2 MALDI mass spectra of CZE fractions (A-F) collected from Figure 4.1B and of the original *E. coli* sample (O).

Table 4.1 Molecular masses of proteins identified from MALDI mass spectra and their identities based on molecular mass matching with *E. coli* proteome databases.

Mass from CZE-MS	Mass from LC-CZE-MS	Mass of Possible Match	Name and Description
	2494.2	2496	NUSA protein (fragment)
2578.9	2580.0	2579	Replication initiator protein (fragment)
2621.8		2621.78	Entericidin B
	2851.8	2852.13	Lysis protein
	2918.1	2918	Hypothetical 2.9 kDa protein (fragment)
	3143.9	3145.77	LEU operon leader peptide
	4138.3	4136.76	Heat-stable enterotoxin 1
	5707.7	5707	Z3951 protein
	5938.2	5936.84	GCV operon transcriptional regulator
	6029.8	6026.89	Hypothetical 6.0 kDa protein in AEFA-PRIC intergenic region
	7158.8	7157.74	50S ribosomal protein L35 (ribosomal protein A)
	7271.6	7271.17	Cold shock-like protein CSPC (CSP-C)
	7333.3	7333.54	Entry exclusion protein B (TRBK protein)
	7725.6	7726.97	YAIA protein
	7739.0	7739.39	ORF70
	7741.8	7741.01	Lysis protein S homolog from lambdoid prophage QIN
7746.3		7745.13	Heme exporter protein D (cytochrome C-type biogenesis protein CCMD)
	7923.2	7923.10	S protein
8224.9		8225	BFPA (fragment)
	8357.0	8359	Hypothetical 8.4 kDa protein
	8450.6	8449.89	Hypothetical 8.5 kDa protein L0053 (ORF2)
	9093.7	9094.54	Hypothetical 9.1 kDa protein in FTSY-NIKA intergenic region
	9113.2	9114.11	YHEA protein
9119.9		9119.37	Phosphocarrier protein HPR (histidine-containing protein)
	9754.1	9753.24	SLT-I B subunit
	9758.1	9755.8	Hypothetical 9.8 kDa protein in DING-GLNQ intergenic region
9767.4		9767.45	SSB24=(GAL4(1-147))-SSB24 transcriptional repressor component
	10342.5	10340.59/ 10344.46	Negative regulator of flagellin synthesis (anti-SIGMA-28 factor)/ hypothetical lipoprotein YBFN
	10528.0	10528.21	Hypothetical 10.5 kDa protein
	10530.3	10529.14	PERC
10534.8		10534.31	Hypothetical protein
	10752.4	10750.25/ 10753.68	Probable SIGMA(54) modulation protein (ORF3) (ORF95)/ hypothetical 10.8 kDa protein
	11828.4	11825	Hypothetical 11.8 kDa protein in IBPA-DGOT intergenic region (precursor)
	11976.5	11976.13	Protein YGIW
	12055.7	12055	Putative transcriptional regulator
15917.8	15917.4	15917.27	F200

4.3.2 Pre-Separation with Preparative LC. The best way to provide a more complete separation of a complex sample is to use orthogonal separation mechanisms. Separation by electrophoresis is based on the differential charge-to-size ratio of molecules. Separation by reverse-phase liquid chromatography, which has been previously employed in the separation of *E. coli* proteins[1], is based on the hydrophobic interaction of proteins with the stationary phase. In general, proteins elute in order of increasing molecular mass. The two separation techniques should therefore complement each other. Even when optimized, however, a typical LC *E. coli* separation with gradient elution takes approximately one hour[2]. Since LC was only to be used as a crude pre-separation method in this work, a fast and simple sample cleanup technique, with large sample loading capacity, was desired. To this end, a laboratory-packed preparative column (0.5 cm x 4.5 cm) of C8 particles was used instead of an analytical column.

The experimental procedure for the LC pre-separation is described above in Section 4.2.5. The entire process of elution and collection of the 19 fractions was completed in less than 5 minutes. MALDI-MS measurements were performed on these LC fractions to determine the presence of proteins. The proteins of interest, with molecular masses over 2000 Da, were detected in the thirteenth to the nineteenth fractions. The MALDI mass spectra of these seven fractions are shown in Figures 4.3A-G respectively. Figure 4.3O presents the mass spectrum of the raw extract prior to LC separation. The small differences in the spectra of Figures 4.2O and 4.3O are due to minor variations associated with the cell extraction and MALDI sample preparation; relative peak intensities in MALDI spectra are very sensitive to the experimental conditions used.

A number of protein peaks are observed in Figures 4.3A-G. Some of these are new proteins that were not observed in Figure 4.3O, such as the 10528 Da protein in fractions A and B. The detection of new proteins indicates a reduction of ion suppression in MALDI analysis. Some proteins are detected in only one fraction, such as the 10753 Da peak in fraction E, and others appear in more than one fraction. Nevertheless, differences in the relative peak intensity are observed for the recurring proteins, for example between the 10530 Da and 15917 Da peaks in fractions A and B, and between the 8450 Da and 10752 Da peaks in fractions D and E. The LC method was therefore

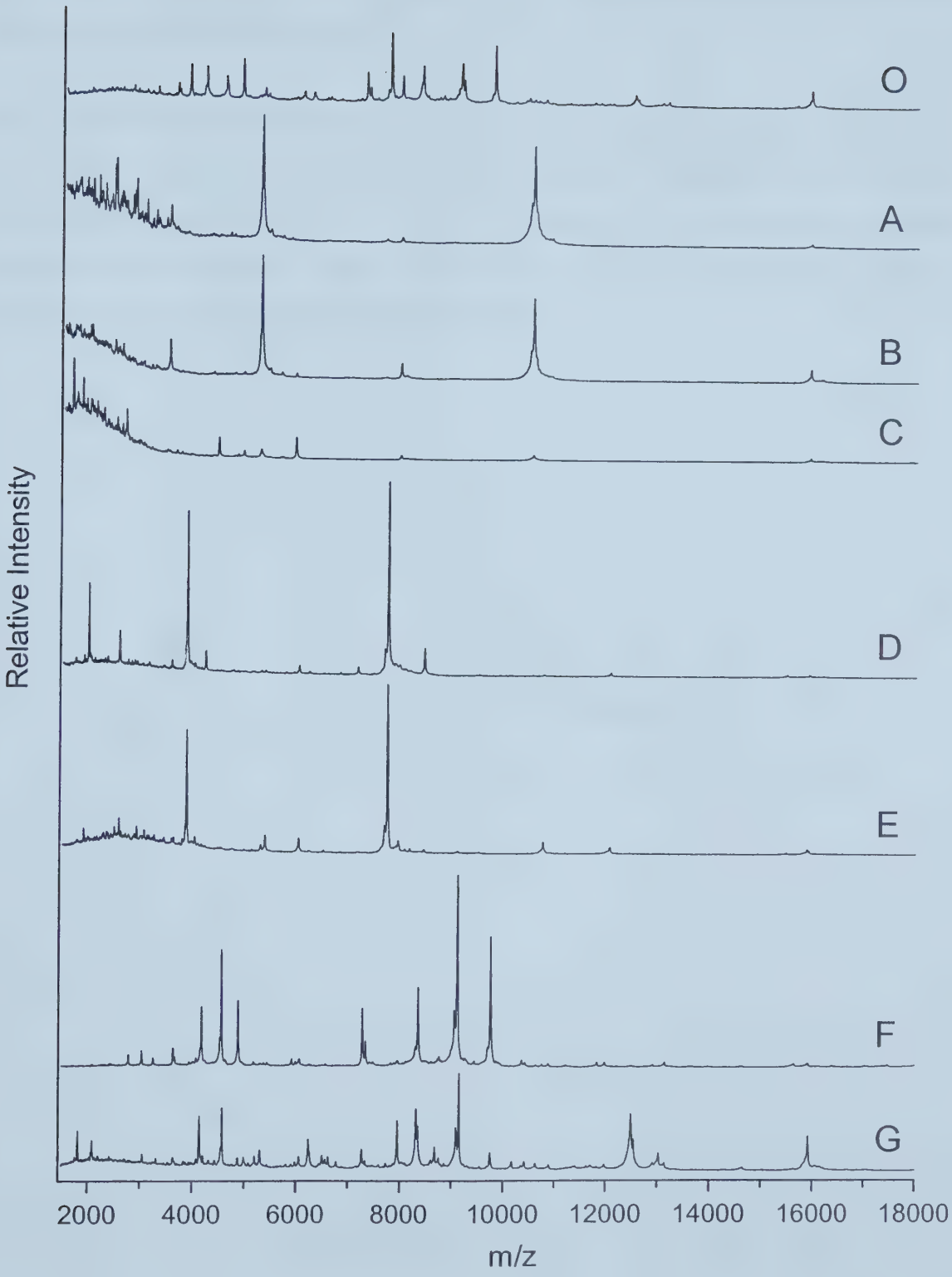

Figure 4.3 MALDI mass spectra of LC fractions (A-G) and of the original *E. coli* sample (O).

successful in achieving a crude separation. The LC pre-separated fractions could then be injected for further separation by CZE.

4.3.3 CZE after LC Pre-Separation. Prior to CZE injection, the seven LC fractions (A to G) were evaporated from 500 μL to *ca.* 5 μL to remove acetonitrile and to concentrate the proteins. Injection times from 1.0 to 10.0 seconds were used for the various fractions depending on the amount of protein present. The electropherograms from the seven LC fractions are shown in the insets of Figures 4.4-4.10.

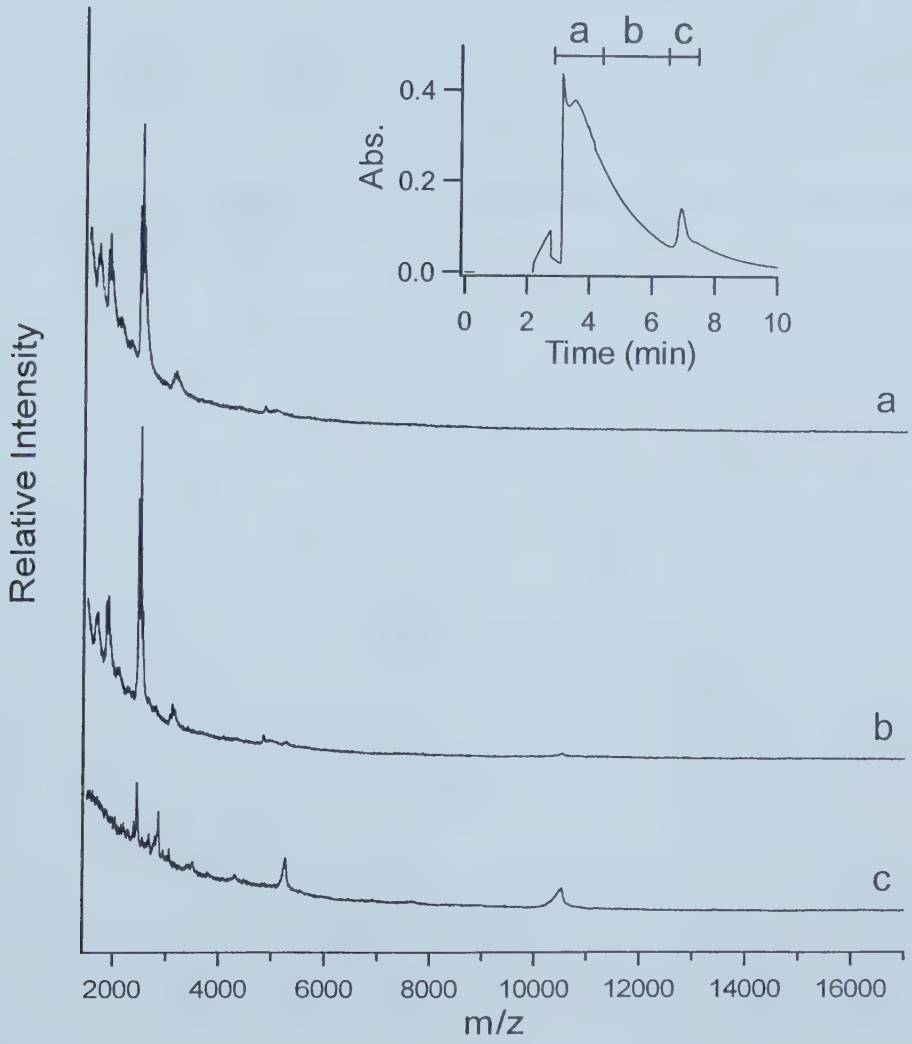

Figure 4.4 Electropherogram (inset) of LC fraction A and MALDI mass spectra of the corresponding CZE fractions.

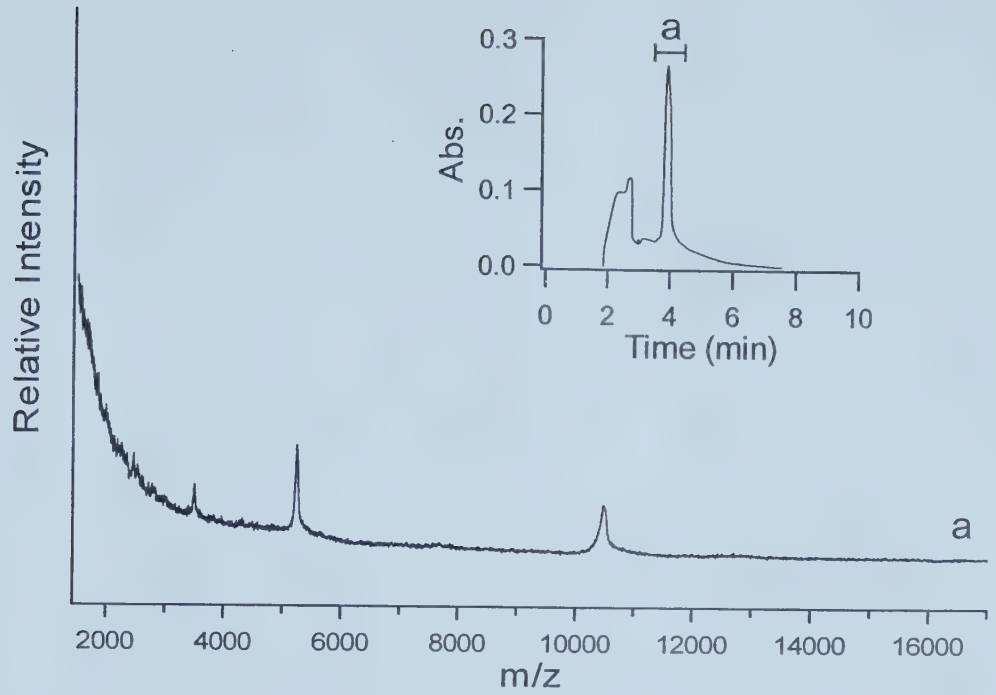

Figure 4.5 Electropherogram (inset) of LC fraction B and MALDI mass spectra of the corresponding CZE fractions.

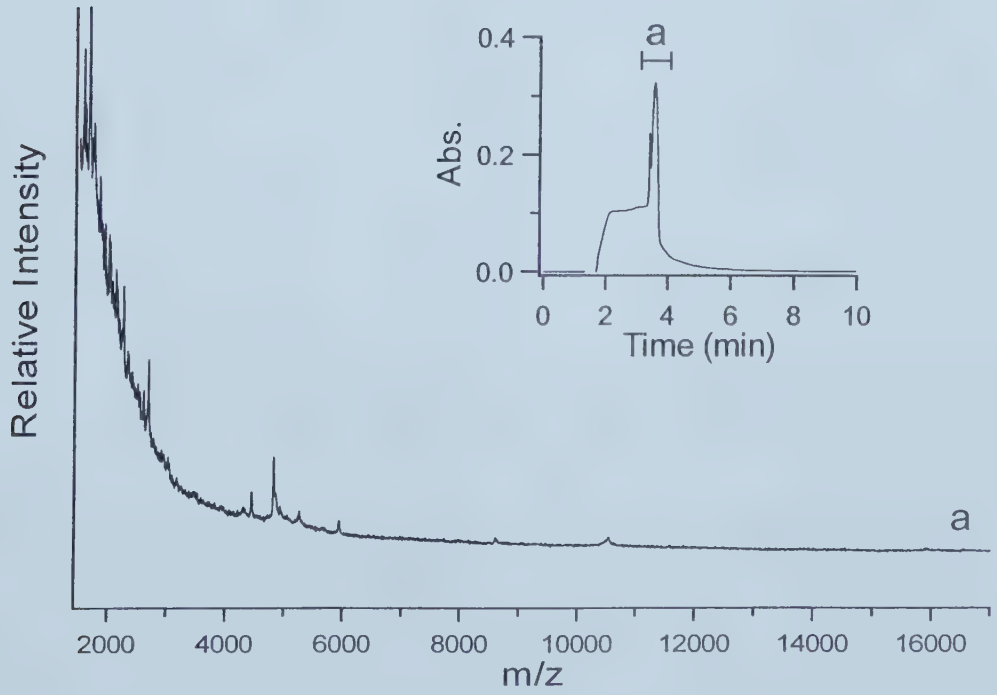

Figure 4.6 Electropherogram (inset) of LC fraction C and MALDI mass spectra of the corresponding CZE fractions.

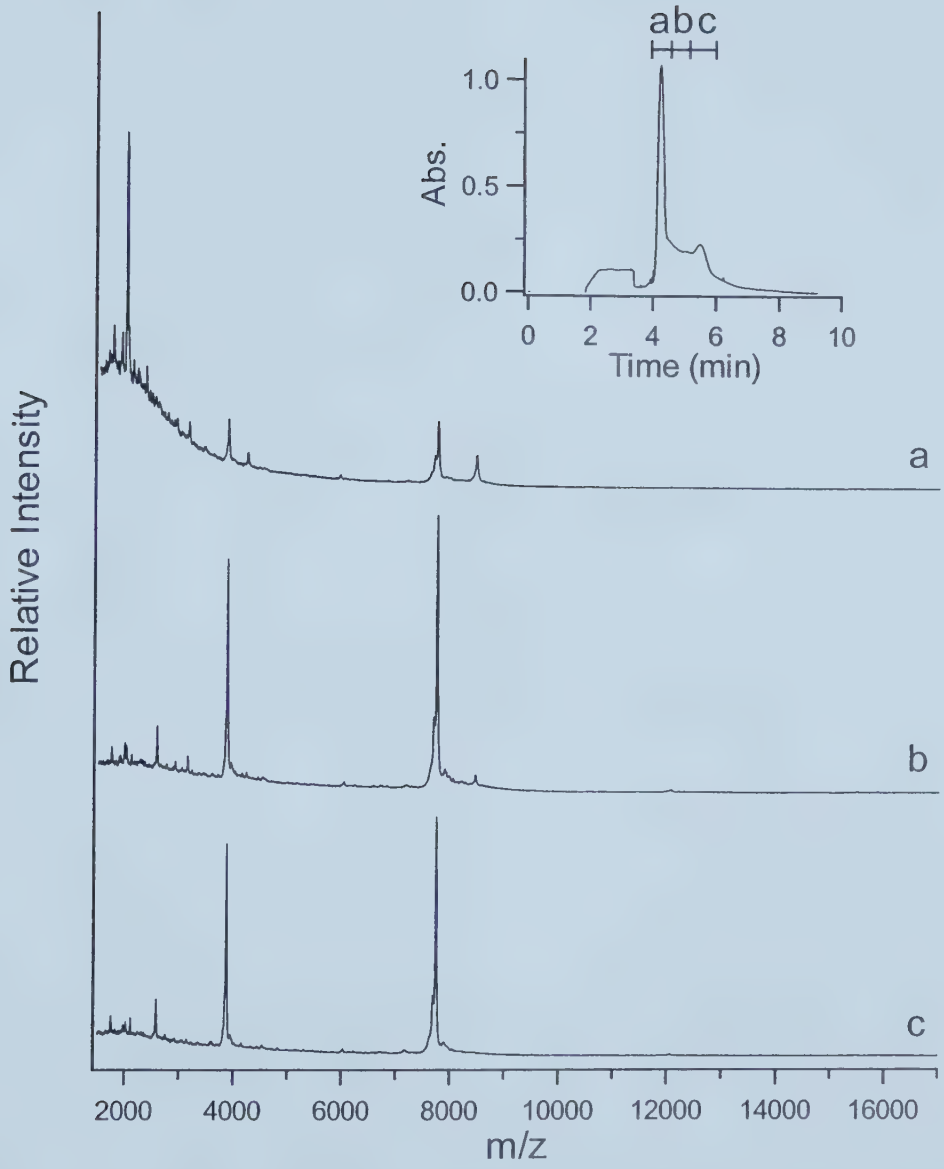

Figure 4.7 Electropherogram (inset) of LC fraction D and MALDI mass spectra of the corresponding CZE fractions.

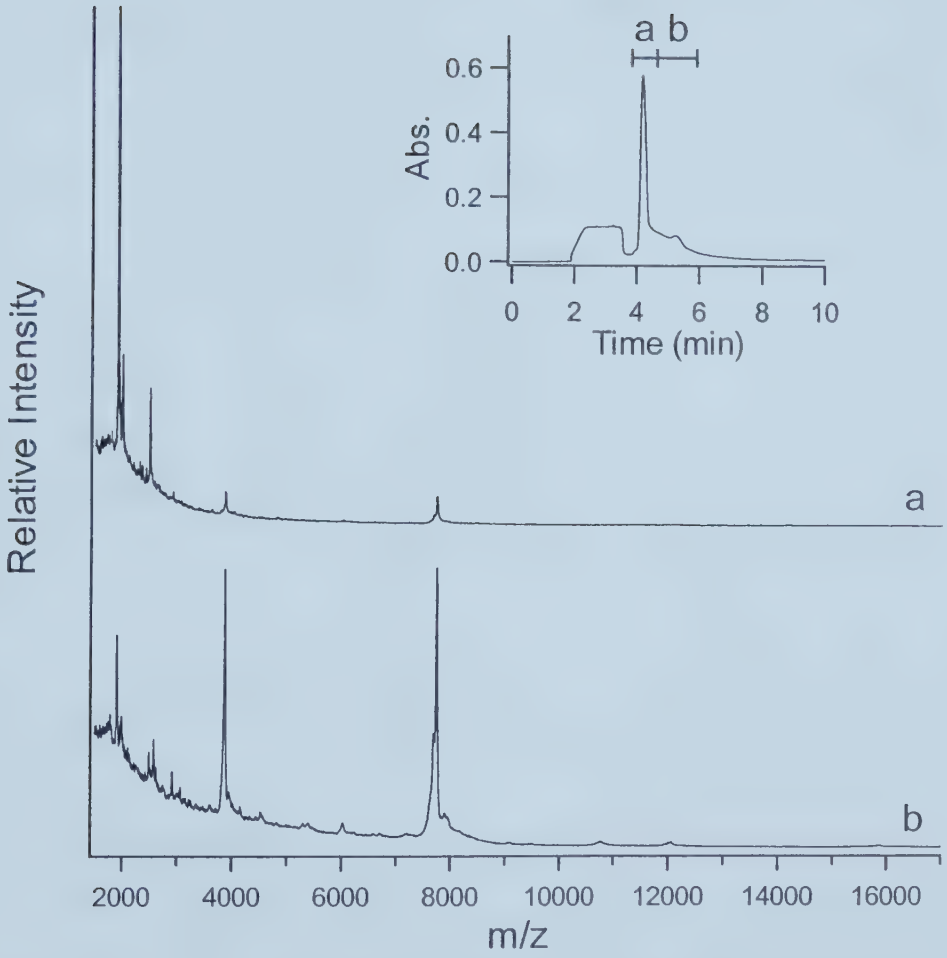

Figure 4.8 Electropherogram (inset) of LC fraction E and MALDI mass spectra of the corresponding CZE fractions.

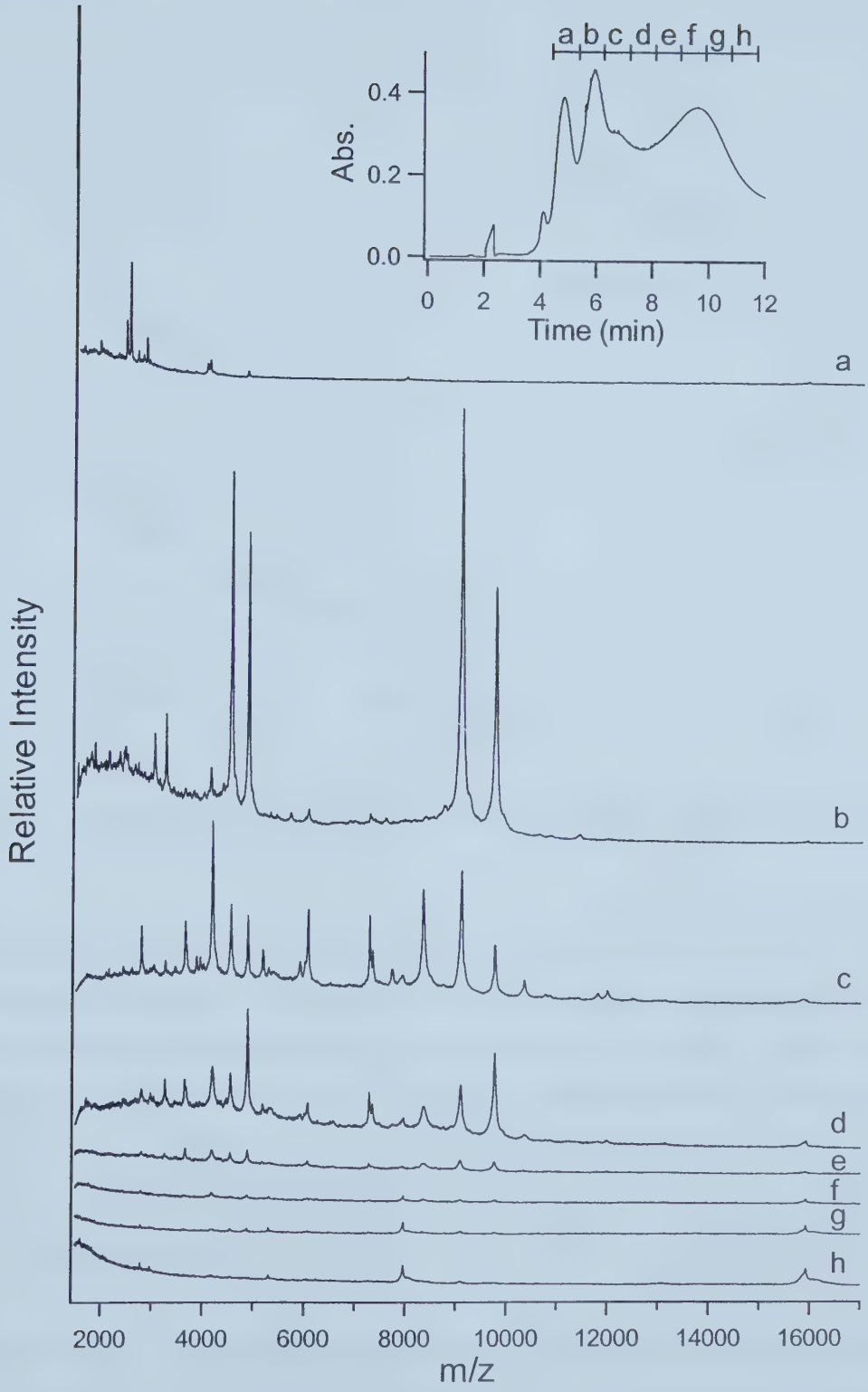

Figure 4.9 Electropherogram (inset) of LC fraction F and MALDI mass spectra of the corresponding CZE fractions.

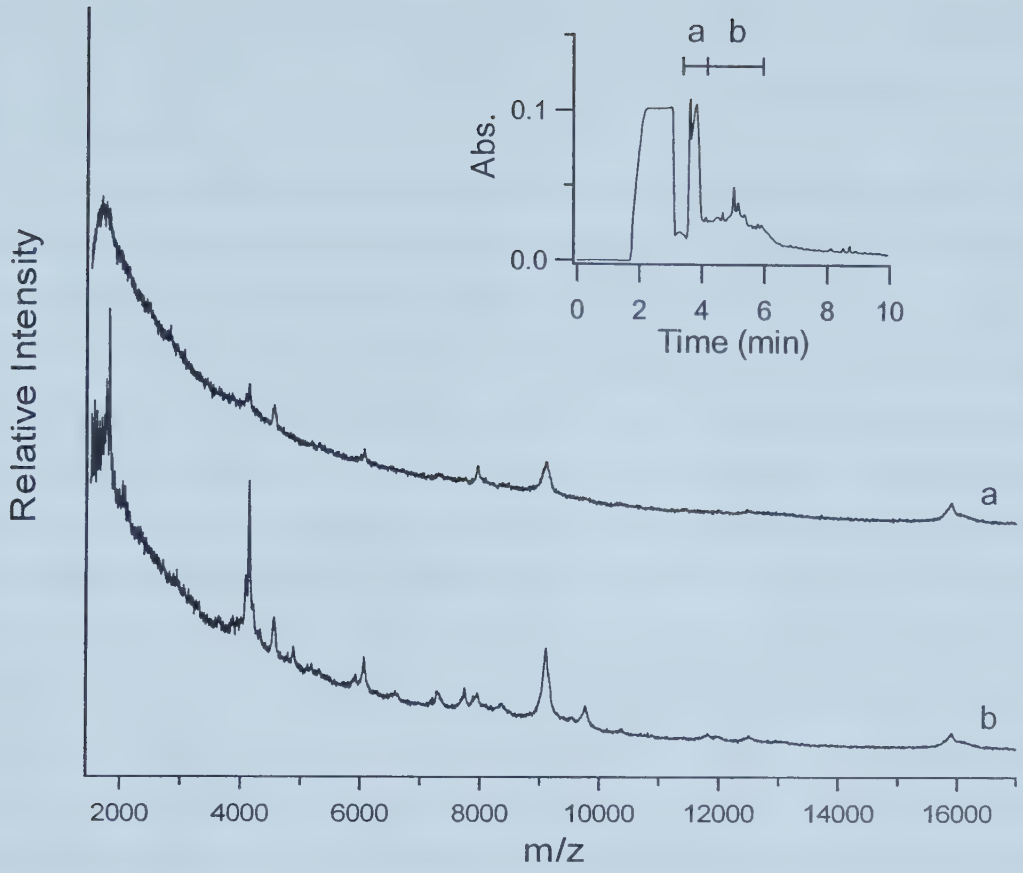

Figure 4.10 Electropherogram (inset) of LC fraction G and MALDI mass spectra of the corresponding CZE fractions.

All of the electropherograms begin with the fronted TFA peak at *ca.* 2 minutes. For long injection times, the TFA peak broadens to a trapezoid-shaped signal as a result of serious injection band broadening. Following the TFA signal, protein peaks are detected in each sample up to 12 minutes. The protein peaks observed are still broad and not fully resolved. Although the sample was pre-separated by LC prior to CZE, presumably a similar degree of electrodispersion and injection band broadening remained. Hence a significant improvement in peak efficiency is not observed.

4.3.4 CZE Fraction Collection and MALDI-MS. MALDI-MS analysis was used to identify the proteins separated by CZE and to evaluate the performance of the CZE separation. First, electrokinetic fraction collection was performed on each CZE separation. The number of fractions collected ranged from one to eight depending on the signals observed in the electropherogram. Then a MALDI mass spectrum was recorded

for each of the CZE fractions collected. These spectra are shown in Figures 4.4-4.10. The time and duration of CZE fractionation are illustrated graphically on the inset electropherograms.

In general, the MALDI mass spectrum of each CZE fraction is unique. In some cases different proteins are present in adjacent fractions. In other cases proteins recur in adjacent fractions, but a shift in the relative intensity is always observed. Overall, more proteins are detected by MALDI after the CZE separation of the LC fractions than could be detected in the LC fractions themselves. This is best illustrated by a comparison of the results obtained in Figure 4.9 to those in Figure 4.3F. The MALDI mass spectrum of LC fraction F (Figure 4.3F) shows a few protein peaks, but many more are detected after CZE fractionation (Figure 4.9). Separation by LC-CZE also results in the detection of more proteins than separation by CZE alone, including some that are not detected in the original *E. coli* extract. The proteins identified from the mass spectra of all the LC-CZE fractions are listed in the second column of Table 4.1. The number of proteins identified after LC-CZE separation is almost four times that detected with CZE separation alone (column 1 of Table 4.1). This indicates that the LC-CZE separation reduces the number of proteins per fraction, thus decreasing analyte interference in MALDI-MS and allowing detection of previously suppressed signals.

A final remark is necessary regarding the intensity of the mass spectra of CZE fractions. Lower signal intensities are observed from the spectra of CZE fractions (panels a, b, c… in Figures 4.4-4.10) compared to those of the LC fractions prior to CZE (Figure 4.3A-G). The MALDI spectra of the LC fractions in Figure 4.3 are obtained without any concentration or dilution. For the MALDI-MS of the CZE fractions, the LC fractions are concentrated 100-fold by evaporation prior to CZE injection. A hydrodynamic injection at 5 kPa is performed for 1.0 to 10.0 seconds, delivering from 3 to 30 nL of sample. After CZE, the separated components are collected into 4 μL of HCl solution. This results in significant dilution, for which even the initial sample concentration cannot fully compensate. Hence weaker MS signals are expected from the CZE fractions. This becomes a problem for low concentration samples, such as LC fraction G. Many signals detected prior to CZE (Figure 4.3G) are lost after CZE fractionation (Figure 4.10), even

when an extended injection time (10.0 seconds) and reduced volume of HCl solution (2 μL) are used.

In addition to dilution, another source of protein loss is the adsorption of proteins onto surfaces. It was found that the MALDI-MS measurement of CZE fractions was best performed immediately after collection. Protein adsorption takes place on surfaces such as the vial, electrode and capillary, presumably via both electrostatic and hydrophobic interactions. Because the volume and concentration of protein in the samples are very low, the effect of protein loss is magnified. However, the adsorption problem associated with protein analysis is not specific to capillary electrophoresis or fraction collection. The strong tendency of hydrophobic proteins to adsorb onto surfaces applies to other techniques, such as HPLC and gel electrophoresis. To make the most of techniques such as CZE that involve the manipulation of very small amounts of protein, improved techniques in protein recovery after separation will have to be investigated.

4.4 Literature Cited

(1) Dai, Y.; Li, L.; Roser, D. C.; Long, S. R. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 73-78.

(2) Dunlop, K. Y.; Li, L. *J. Chromatogr. A* **2001**, *925*, 123-132.

(3) Dalluge, J. J. *Fres. J. Anal. Chem.* **2000**, *366*, 701-711 and references therein.

(4) O'Connor, C. D.; Adams, P.; Alefounder, P.; Farris, M.; Kinsella, N.; Li, Y.; Payot, S.; Skipp, P. *Electrophoresis* **2000**, *21*, 1178-1186 and references therein.

(5) Tang, Q.; Harratta, A. K.; Lee, C. S. *Anal. Chem.* **1997**, *69*, 3177-3182.

(6) Jensen, P. K.; Pasa-Tolic, L.; Anderson, G. A.; Horner, J. A.; Lipton, M. S.; Bruce, J. E.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2076-2084.

(7) Jensen, P. K.; Pasa-Tolic, L.; Peden, K. K.; Martinovic, S.; Lipton, M. S.; Anderson, G. A.; Tolic, N.; Wong, K.-K.; Smith, R. D. *Electrophoresis* **2000**, *21*, 1372-1380.

(8) Zheng, J.; Wang, Z.; Li, L. In *49th American Society of Mass Spectrometry Conference*: Chicago, IL, 2001.

(9) Melanson, J. E.; Baryla, N. E.; Lucy, C. A. *Anal. Chem.* **2000**, *72*, 4110-4114.

(10) Figeys, D.; Aebersol, R. *Electrophoresis* **1998**, *19*, 885-892 and references therein.

(11) Bergman, A. C.; Bergman, G. *J. Prot. Chem.* **1997**, *16*, 421-423.

(12) Keller, B. O., *Ph.D. thesis*; The University of Alberta: Edmonton, 2001.

(13) Walker, W. L.; Chiu, R. W.; Monnig, C. A.; Wilkins, C. L. *Anal. Chem.* **1995**, *67*, 4197-4204.

(14) Whittal, R. M.; Li, L. *Anal. Chem.* **1995**, *67*, 1950-1954.

(15) Whittal, R. M.; Russon, L. M.; Weinberger, S. R.; Li, L. *Anal. Chem.* **1997**, *69*, 2147-2153.

(16) Roberts, G. O.; Rhodes, P. H.; Snyder, R. S. *J. Chromatogr.* **1989**, *480*, 35-67.

(17) Yeung, K. K.-C.; Kiceniuk, A. G.; Li, L. *J. Chromatogr. A* **2001**, in press.

(18) Lucy, C. A.; Yeung, K. K.-C.; Chen, D. D. Y.; Peng, X. *LC•GC* **1998**, *16*, 26-32.

Chapter 5

Conclusions and Future Work

This thesis explores the combination of capillary electrophoresis with MALDI mass spectrometry for proteome analysis. In Chapter 2 it is shown that micro fraction collection in CE is a simple yet effective method of coupling CE with MALDI MS for the analysis of whole proteins. The timing of the fraction collection is well controlled using electrokinetic fractionation, and the collection of multiple consecutive fractions in one CE separation within a short time window is demonstrated. The choice of electrolyte for CE fraction collection is critical in order to minimize interference during MS analysis. It is shown that the use of dilute acid for fraction collection does not alter the CE separation efficiency and fraction collection accuracy while maintaining low MS detection limits. It is also demonstrated that the use of capillaries treated with the two-tailed surfactant DDAB prevents protein adsorption and provides a fast EOF for short separations. Unlike other surfactants, DDAB forms a semi-permanent wall coating, and thus is compatible with subsequent MALDI MS measurements.

To further develop this CE/MALDI MS method into a powerful tool for proteomics applications, a method of gaining more information about the proteins analysed is necessary. Chapter 3 discusses the use of peptide mapping for the identification of CE-separated proteins. On-target digestion methods proved to be the most effective, giving up to 79% sequence coverage with as little as 50 fmol of cytochrome *c*. The AnchorChip™ digestion results did not meet our expectations, but further work on improving the matrix crystallization will likely improve the results. At lower analyte concentrations, PEG contamination becomes a major obstacle. In digests of 20 fmol of cytochrome *c*, PEG overwhelms the analyte peptide peaks in the mass spectrum. The source of the contamination must be found and eliminated if peptide mapping is to be extended to lower concentration samples. If the PEG contamination cannot be eliminated, a sample preparation method that concentrates the proteins while

eliminating the PEG will have to be used prior to digestion. An affinity-based preconcentration step is one such possibility.

Chapter 4 demonstrates the CE separation of a complex protein mixture from *E. coli* cell extracts. Only limited CE resolution can be achieved with so many proteins analysed in such a short time, so LC is employed as a pre-separation before CE to improve the resolution. LC/CE reduces analyte interference in MALDI-MS measurement and increases the number of proteins identified by a factor of almost four. Some important issues related to the development of LC/CE/MALDI MS as a practical tool for proteome analysis are also identified in this chapter. Future work will focus on improving separation efficiencies in both LC and CE for complex protein mixture analysis, developing hardware for better LC to CE to MALDI MS interfaces, and implementing enzymatic digestion of final fractions for protein identification based on peptide mass mapping and fragment ion spectra of peptides.

CE/MALDI MS holds great promise as a powerful technique for proteome analysis. The high efficiency, fast analysis time, low sample loading and multiplex capability of CE make it an ideal match for the high-throughput demands of contemporary proteomic research, especially when combined with the high sensitivity of MALDI MS. As it is further refined in the future, CE/MALDI MS could become a widely used tool along with other multidimensional microseparation techniques in analytical and clinical laboratories.